In conclusion, brain ChE was not altered at either dose of CPFO; thus, the 0.01 mg/kg/day CPFO dose was considered a no-observed-effect-level (NOEL) for ChE inhibition across all tissues in both age groups with repeated exposure.

4. Discussion

This comparative cholinesterase study was designed to examine whether there were age-related differences in sensitivity to ChE inhibition following CPF exposure. The effects of acute CPF exposure in PND 11 pups compared with adults showed that there were dose-dependent differences in sensitivity at high doses (i.e., 5 mg/kg in corn oil), where pups were more sensitive than adults. At lower dose levels, pups and adults showed significant plasma and RBC or brain ChE inhibition at the same dose levels. When using rat milk as an alternate vehicle to simulate lactational exposures, ChE inhibition in PND 11 pups was similar to levels achieved when administering CPF in corn oil. In adults, acute CPF exposure by gavage (corn oil vehicle) or using a alternate dosing scenario (i.e., 12-h dietary exposure in adult females) showed that dose rate affects the relative magnitude of ChE inhibition across tissues, because ChE inhibition was greater in RBC, but less in brain, with dietary exposures. With 11 daily exposures to CPF (corn oil vehicle), both pups and adults showed similar sensitivity to ChE inhibition across tissues. Brain ChE was not inhibited in either adults or pups at any dose of CPFO tested, despite similar sensitivity at ages to CPFO inhibition of RBC and plasma ChE. Together, these data indicate that young animals are not more sensitive than adults to CPF- or CPFO-induced ChE inhibition across the lower portion of the dose–response curves.

PND 11 pups were selected as the comparison group for adult females in this study based on EPA guidance for comparative cholinesterase studies. Across species, brain development follows predetermined developmental patterns; however, the timing of these developmental stages relative to birth varies. Consequently, the degree of functional maturity of the nervous system also varies at birth. There is ample data to support the concept that compared to humans, rats are altricial (i.e., born less mature) with respect to neurodevelopment. Using morphometric measurements and neurogenesis in different brain regions, Bayer et al. (1993) concluded that the full-term human brain at birth was approximately equal to a PND 14–21 rat brain. Similar conclusions also have been expressed by others (Vidair, 2004; Clancy et al., 2007). Data suggest that the blood–brain barrier (BBB) in humans is more developed at birth than the BBB in neonatal rats (Adinolfi and Haddad, 1977; Bonati et al., 1981). Thus, PND 11 rat pups represent an appropriately conservative model to examine potential effects in human infants.

When examining control animals from the acute phase of the comparative ChE study, adults had greater brain and plasma ChE activity, whereas pups had greater RBC ChE activity. In the repeat-dose study, when pups were euthanized on PND 21, brain ChE activity in the control group was approximately 80% higher than levels in PND 11 control pups. This was consistent with previous reports (Moser et al., 1998; Timchalk et al., 2006) that have demonstrated increases in brain ChE as rats mature. There was no apparent maturational pattern for RBC and plasma ChE activity, which were approximately similar at both PND 11 and 21. Timchalk et al. (2006) reported that RBC ChE enzyme activity increased between PND 5 and 12, then decreased slightly on PND 17, whereas plasma ChE activity was relatively stable across ages.

With acute exposures in corn oil, the relative sensitivity of adults compared with pups was dose dependent. At high dose levels, PND 11 pups were more sensitive to CPF-induced ChE inhibition because 5 mg/kg CPF induced similar levels of plasma, RBC and brain ChE inhibition as 10 mg/kg in adults. However, at lower doses of CPF, the dose–response curves for adults and immature rats intersected (see Figure S-21; Supplemental data 1), such that 2 mg/kg did not cause significant brain ChE inhibition in either adults or pups, but caused significant RBC and plasma ChE in both age groups. As reported previously (US EPA, 2011), significant RBC and plasma ChE inhibition occurred at lower dose levels than brain ChE inhibition in both PND 11 pups and adults. The NOEL for ChE inhibition across all tissues (0.5 mg/kg) was the same for both adults and pups. There was no significant difference in sensitivity to ChE inhibition between male and female PND 11 pups, consistent with other reports in preweanling animals (e.g., Moser and Padilla, 1998; Moser et al., 1998). Thus, at lower doses, adult female rats and PND 11 rat pups exhibited similar sensitivity across all tissues, although pups had higher blood levels of CPF at all dose levels.

The enhanced sensitivity of pups to acute CPF exposure at higher dose levels has been reported previously (e.g., Moser et al., 1998 at doses >5 mg/kg) and was partially attributed to the lower metabolic capacity in younger animals (Timchalk et al., 2006). There is evidence that pups metabolize high doses of CPF more slowly than adults. When examining [TCP]/[CPF] ratios across studies, which is an indicator of metabolic capacity of an animal, it appears that this ratio is ~50 in PND 5 pups at 6 h after dosing 1 mg/kg CPF in corn oil (Marty et al., 2007), 102–170 in PND 11 pups at 6 h after dosing with 0.5–2 mg/kg CPF in corn oil and ~449–810 in adult females at 8 h after dosing 0.5–2 mg/kg CPF in corn oil. These data show greater metabolic capacity was present in adults; however, TCP formation was favored in the older age groups as blood TCP concentrations exceeded parent CPF by a factor of 100-fold or greater by PND 11, consistent with the findings of Timchalk et al. (2006).

The production of CPFO depends on the rate of hepatic activation of CPF and inactivation of CPFO by cytochrome P450 monooxygenases (Ma and Chambers, 1994; Sultatos, 1994; Sultatos et al., 1984). CPF is extensively metabolized into water soluble metabolites, which prevents the accumulation of CPF or its metabolites (US EPA, 2011). Other pathways involved in CPFO inactivation include interactions with esterases other than AChE (e.g., butyrylcholinesterases, which is hypothesized to scavenge CPFO to prevents its interaction with peripheral target sites; Maxwell 1992a,b; US EPA, 2011) or binding to B-esterases (e.g., carboxylesterases), both of which decrease the amount of CPFO available to interact with the target site (AChE). In addition, hydrolysis of the oxon by A-esterases (i.e., PON-1; CPF-oxonase; Behnke and Murphy, 1975; Costa et al., 1990) leads to the formation of TCP and diethylphosphate, which do not inhibit AChE. Data indicate that these detoxification pathways continue to mature postnatally in rats and the maturation of these systems parallels decreases in sensitivity to high-dose, acute CPF exposure (Mortensen et al., 1996; Atterberry et al., 1997; Maxwell, 1992a,b; Chand et al., 1997; Morgan et al., 1994). Moser et al. (1998) showed that preweanling rats have lower levels of both liver and plasma carboxylesterases and A-esterase activity than adults, which correlates with the gradual decrease in sensitivity as rats mature. However, in humans, available data indicate that liver carboxylesterase activity does not differ between infants and adults as activity appears to change relatively little during postnatal maturation (Pope et al., 2005). Furthermore, Smith et al. (2011) found no age-related differences in CPF metabolism *in vitro* using hepatic microsomes isolated from humans at 13 days to 75 years old, whereas age-dependent increases in CPFO esterase metabolism in human plasma (3 days to 46 years) were reported.

Levels of ChE inhibition following acute ChE exposure in the current study were generally consistent with previously published studies in immature animals, although this study included multiple dose levels at the lower portion of the dose response curve (i.e., <1 mg/kg). In the study by Timchalk et al. (2006), RBC and plasma inhibition were seen in PND 5 and 12 pups at 1 mg/kg

CPF in corn oil. This is consistent with the current study, where inhibition was seen in these tissues at 2 mg/kg, but was not seen at 0.5 mg/kg in PND 11 pups. Zheng et al. (2000) reported a decrease in plasma and RBC ChE at doses of 0.45 and 1.5 mg/kg CPF, respectively, in PND 7 pups. The reason for this difference in plasma ChE inhibition may be related to differences in study designs, sampling times, or ages of the pups from which ChE activity was measured. In the current study, there were no effects on brain ChE activity at 2 mg/kg on PND 11, which was consistent with Timchalk et al. (2006), who reported no effects on brain ChE activity at 1 mg/kg in PND 12 pups. Overall, when considering dose and pup age, the levels of ChE inhibition across tissues in this study were consistent with the existing scientific literature.

The variability in ChE measurements in the current study were consistent with variability reported in other studies, giving the current study a similar level of sensitivity to previous work. Coefficients of variation (CVs) for control ChE values appear in Supplemental data 3. In the acute dose–response studies (n = 8/sex/dose for pups or n = 8/dose for adult females), RBC CVs ranged from 6.5% to 23.2%, brain CVs ranged from 2.3% to 18.1% and plasma CVs ranged from 7.5% to 37.5%. In the repeat dose studies (n = 8/sex/dose for pups or n = 8/dose for adult females), RBC CVs ranged from 5.6% to 35.0%, brain CVs ranged from 3.5% to 3.8% and plasma CVs ranged from 12.3% to 33.0%. These results were consistent with expectations as brain ChE activity was the least variable of the tissues measured, whereas there was more variability in plasma ChE, which contains mixed activity (i.e., butyryl- and acetyl-cholinesterase). A CV comparison for treated animals was not included as variance was expected to be higher in treated animals due to inter-animal differences, including differences in absorption, distribution, metabolism, and excretion, individual differences in response to treatment (particularly in steeper portions of the dose–response curve), slight differences in dose delivered, etc. The similarity in CVs to other published studies shows that these assays were reasonably sensitive to detect changes in ChE activity when such changes were present.

When using rat milk as an alternative vehicle in PND 11 pups to simulate lactational exposures, ChE inhibition was similar to levels achieved when administering CPF in corn oil. At the time-of-peak inhibition, blood CPF and CPFO levels, as well as the magnitude of ChE inhibition across tissues, were similar with both milk and corn oil vehicles (Fig. 1A and B). This was unexpected as kinetic data for blood CPF and blood TCP in PND 5 pups (1 mg/kg CPF in corn oil or in milk) showed a similar time to maximal concentration (C_{max}) for both oil and milk vehicles with a notable increase in blood CPF C_{max} in pups dosed with corn oil (Marty et al., 2007). Based on the established PBPK/PD model for CPF in immature rat pups (Timchalk et al., 2002, 2006), the peak for pup blood levels of CPF after administration in rat milk was 5–7 h. Given the slow recovery of ChE activity, the ChE inhibition was comparable even with different time points examined (6 vs. 8 h post-dosing). In their recent assessment, the US EPA determined that RBC ChE inhibition in PND 11 pups exposed acutely to CPF in milk had the lowest oral point of departure in the CPF database (US EPA, 2011).

When using an acute 12-h dietary exposure in adults to simulate CPF exposures in the diet over a day, dose rate apparently impacted the relative magnitude of tissue ChE inhibition. The slower dose rate likely allowed more time for detoxification pathways, such that less CPF was available to interact with brain ChE. Data from these studies have shown that CPFO at <10 mg/kg did not inhibit brain ChE activity in adult females. Therefore, it is possible that the slower dose rate allowed greater opportunity for P450 metabolism and/or interaction of CPFO with carboxylesterases or other ChEs (e.g., butyrylcholinesterase) so that less CPFO was available to interact with brain ChE. Timchalk et al. (2006) reported that differences in tissue dosimetry (higher oxon AUC in blood relative to brain) contribute to enhanced sensitivity of blood relative to brain ChE. This seems plausible as RBC ChE showed greater inhibition with dietary CPF dosing, although plasma ChE inhibition was the same with both gavage and dietary treatment.

In the current study, there were no signs of cholinergic toxicity detected in either the acute study (clinical observations at ⩽5 mg/kg in pups or ⩽10 mg/kg in adults) or the repeat-dose study (clinical observations and FOB with motor activity at ⩽3.5 mg/kg/day in both age groups), despite significant brain ChE inhibition (~53–58% in the acute study and 59–69% in the repeat dose study). Moser (2000) reported that PND 17 female rats had a decrease in tail-pinch response at 6.5 h post-dosing with 4 mg/kg CPF, whereas males were not affected at this dose level. At a higher CPF dose (i.e.,10 mg/kg) than those given to pups in the current study, Moser (2000) observed alterations in numerous neurobehavioral endpoints in both male and female PND 17 pups including altered gait/ataxia, decreased arousal state, tail-pinch response (males), tremors, smacking (males) and lacrimation (females). In adults, decreased motor activity (total counts) was the most sensitive endpoint with decreases in males noted at 3.5 h post-dosing with 10 mg/kg, whereas both males and females were affected at 50 mg/kg, along with other cholinergic signs of toxicity (Moser, 2000). In a separate study (Moser et al., 1998), PND 17 female rats with a 50–60% decrease in brain ChE activity following acute exposure to 5 mg/kg CPF showed decreased open-field arousal, and in adult males, there was a correlation between a 40–50% decrease in brain ChE activity and decreased motor activity at 20 mg/kg CPF. Numerous factors may have contributed to differences in neurobehavioral effects in the previous studies by Moser et al. and the current study, including differences in dose levels (generally lower in the present study), developmental stage, neurobehavioral methods, or possibly receptor down regulation, which has been proposed to account for recovery of neurobehavioral performance (Bignami et al., 1975; Nostrandt et al., 1997; Moser and Padilla, 1998). Perhaps the gender-related differences in sensitivity to neurobehavioral effects in the Moser studies, in the presence of similar brain ChE inhibition in both sexes, indicate that these effects occurred near the threshold for neurobehavioral alterations, which could vary slightly across studies.

With repeated exposures, adults and PND 11 pups were similar in sensitivity to ChE inhibition by CPF. Significant brain ChE inhibition was seen in both adults and pups at 1.0 mg/kg/day CPF, whereas significant plasma and RBC inhibition occurred in both age groups at 0.5 mg/kg/day. In 2008, the US EPA Science Advisory Panel (2008) hypothesized that young animals might be less sensitive to repeated CPF exposure due to decreased levels of enzymes converting CPF to CPFO and/or a more rapid increase in AChE activity in tissues of young animals, likely due to increased rates of protein synthesis (e.g., Chakraborti et al., 1993; Liu et al., 1999). The current study verified that pups achieve higher blood levels of CPF for a given dose, presumably due to slower metabolism to TCP. Based on administered dose, these immature animals showed similar sensitivity to CPF-induced ChE inhibition as adults; however, based on blood levels, pups showed lower sensitivity to CPF-induced ChE inhibition. A recent physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) model using human CYP-specific kinetic parameters and age-based differences in hepatic CYP content predicted that 1 year-olds would be less sensitive than 19-year olds to CPF-induced butyryl- and acetyl-ChE inhibition, although age-related differences in PON-1 levels and microsomal liver content are needed to refine the model (Foxenberg et al., 2011).

In adult females, there was significant inhibition of RBC ChE activity at 0.1 mg/kg/day in the repeat-dose study, which was deemed incidental as plasma ChE was not significantly altered at this dose level. Plasma ChE has similar or greater sensitivity than

RBC to inhibition by CPF (Lotti, 1995; Garabrant et al., 2009; US EPA, 2008). Plasma ChE, which is comprised of half butyrylcholinesterase and half AChE (Timchalk et al., 2006), is more readily inhibited because butyrylcholinesterase is more sensitive to inhibition by CPFO than AChE (Amitai et al., 1998; Kousba et al., 2003; Timchalk et al., 2002). Across the current data sets, there were a few occurrences when samples collected at the same time showed RBC ChE inhibition that exceeded plasma ChE inhibition; however, this was in a minority of cases and when it occurred, RBC and plasma ChE samples were similar. Thus, the significant inhibition of RBC ChE in adult females at 0.1 mg/kg/day, which occurred in the absence of significant plasma ChE inhibition, was deemed spurious.

Overall, the dose–response for ChE inhibition with repeated CPF exposure in the current study was consistent with three previous studies examining CPF-induced ChE inhibition in rats. In a 90-day repeat-dose dietary CPF study with F344/DuCrl rats, Szabo et al. (1988) reported a significant decrease in RBC and brain cholinesterase activities at $\geqslant 1$ and $\geqslant 5$ mg/kg/day, respectively. In the adult 28-day dietary immunotoxicity study in CD rats (Boverhof, personal communication), significant RBC and brain cholinesterase inhibition were seen at similar dose levels ($\geqslant 0.4$ and $\geqslant 2$ mg/kg/day CPF, respectively) to the current study. RBC ChE inhibition at 0.4 mg/kg/day CPF was somewhat greater in the immunotoxicity study (53.7% of control compared with 80.5% at 0.5 mg/kg/day in the current study), which may have been related to the extended dosing period in the immunotoxicity study (11 days in the current study vs. 28 days in the immunotoxicity study) or the difference in exposure routes (oral gavage in corn oil in the current study vs. dietary in the immunotoxicity study). In a study by Carr and Nail (2008), ChE inhibition was measured in multiple areas of the brain in rat pups dosed daily by gavage from PND 10–16 with 5 mg/kg/day CPF in corn oil. Brain ChE inhibition ranged from 54% (cerebellum and medulla) to 64% (forebrain) with this dosing paradigm compared with whole brain ChE inhibition that ranged from 59% to 68% at 3.5 mg/kg/day from PND 11–21 in the current study. These results show comparable brain ChE inhibition despite slight differences in dose and exposure duration.

With CPFO exposure, brain ChE was not inhibited in either adults or pups at any dose level tested, despite similar sensitivity at both ages to CPFO inhibition of RBC and plasma ChE. These data indicate a lack of systemic bioavailability of CPFO to peripheral tissues (Bartels et al., 2011) and suggest that exposure to CPFO is less toxic to brain ChE than exposure to parent CPF. In preliminary studies, doses $\leqslant 10$ mg/kg CPFO did not inhibit brain ChE activity in adult female rats (Table 2). This finding differs from Betancourt and Carr (2004) who reported ~50–60% decreases in brain ChE in newborn rats (PND 1–6) exposed daily to 0.25 or 0.35 mg/kg/day CPFO via oral gavage. These results may differ from the current study because the pups were younger at the time of exposure; therefore, an incomplete blood–brain barrier and/or slower detoxification pathways for CPFO may have contributed to brain ChE inhibition. However, this PND 1–10 age range has generally been considered to be physiologically more consistent to human fetuses *in utero* (US EPA, 2011) and therefore, would not be relevant for evaluation of neonatal human exposures.

In conclusion, both the acute and repeated-dose data indicate that young animals are not more sensitive than adults to CPF or CPFO over the lower portion of the dose response curves. This conclusion has been confirmed subsequently using Benchmark Dose Modeling (Reiss et al., 2012). Thus, with low-level, environmentally relevant exposures, higher sensitivity of young animals to ChE inhibition would be unlikely. Furthermore, there is no indication of altered brain ChE activity in either pups or adults following 11 daily exposures to <0.5 mg/kg/day CPFO, the proximate toxicant with CPF exposure. These data suggest that there should be little, if any, concern for CPFO-mediated brain ChE inhibition at environmentally relevant exposure levels.

Conflict of interest

The authors of this article, with the exception of M.J. Beck, are employed by The Dow Chemical Company or Dow AgroSciences, LLC, which produce chlorpyrifos and funded this study.

Acknowledgments

The authors wish to acknowledge the US EPA for advice on dose selection, sample collection times and other aspects of this study that were decided prior to initiation of the definitive study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.yrtph.2012.03.015.

References

Adinolfi, M., Haddad, S.A., 1977. Levels of plasma proteins in human and rat fetal CSF and the development of the blood–CSF barrier. Neuropadiatrie 8, 345–353.
Amitai, G., Moorad, D., Adani, R., Doctor, B.P., 1998. Inhibition of acetylcholinesterase and butyrylcholinesterase by chlorpyrifos–oxon. Biochem. Pharmacol. 56, 293–299.
Atterberry, T.T., Burnett, W.T., Chambers, J.E., 1997. Age-related differences in parathion and chlorpyrifos toxicity in male rats: target and nontarget esterase sensitivity and cytochrome P450-mediated metabolism. Toxicol. Appl. Pharmacol. 147, 411–418.
Bartels, M., Marty, M.S., Hotchkiss, J.A., Juberg, D.R., 2011. Imparct of non-linear pharmacokinetics and metabolism of chlorpyrifos on biological response in the rat. Abstract # 2279. 2011 Itinerary Planner. Society of Toxicology, Washington, DC.
Bayer, S.A., Altman, J., Russo, R.J., Zhang, X., 1993. Timetables of neurogenesis in the human brain based on experimentally determined patterns in the rat. Neurotoxicology 14, 83–144.
Behnke, G.M., Murphy, S.D., 1975. The influence of age on the toxicity and metabolism of methyl parathion and parathion in male and female rats. Toxicol. Appl. Pharmacol. 31, 254–269.
Betancourt, A.M., Carr, R.L., 2004. The effect of chlorpyrifos and chlorpyrifos–oxon on brain cholinesterase, muscarinic receptor binding, and neurotrophin levels in rats following early postnatal exposure. *Toxicol.* Sci. 77, 63–71.
Bignami, G., Rosic, N., Michalek, H., Milosevic, M., Gatti, G.L., 1975. Behavioral toxicity of anticholinesterase agents: methodological, neurochemical, and neuropsychological aspects. In: Weiss, B., Laties, V.G. (Eds.), Behavioral Toxicology. Plenum, New York, pp. 155–216.
Bonati, M., Latini, R., Marra, G., Assael, B.M., Parini, R., 1981. Theophylline distribution in the premature neonate. Dev. Pharmacol. Ther. 3, 65–73.
Brzak, K.A., Harms, D.W., Bartels, M.J., Nolan, R.J., 1998. Determination of chlorpyrifos, chlorpyrifos oxon, and 3,5,6-trichloro-2-pyridinol in rat and human blood. J. Anal. Toxicol. 22, 203–210.
Carr, R.L., Nail, C.A., 2008. Effect of different administration paradigms on cholinesterase inhibition following repeated chlorpyrifos exposures in late prewenling rats. Toxicol. Sci. 106, 186–192.
Chakraborti, T.K., Farrar, J.D., Pope, C.N., 1993. Comparative neurochemical and neurobehavioral effects of repeated chlorpyrifos exposures in young and adult rats. Pharmacol. Biochem. Behav. 46, 219–224.
Chand, S.M., Mortensen, S.R., Moser, V.C., Padilla, S., 1997. Tissue-specific effects of chlorpyrifos on carboxylesterase and cholinesterase activity in adult rats: an *in vitro* and *in vivo* comparison. Fundam. Appl. Toxicol. 38, 148–157.
Clancy, B., Kersh, B., Hyde, J., Darlington, R.B., Anand, K.J.S., Finlay, B.L., 2007. Web-based method for translating neurodevelopment from laboratory species to humans. Neuroinformatics 5, 79–94.
Costa, L.G., McDonald, B.E., Murphy, S.D., Omenn, G.S., Richter, R.J., Motulsky, S.G., Furlong, C.E., 1990. Serum paraoxonase and its influence on paraoxon and chlorpyrifos–oxon toxicity in rats. Toxicol. Appl. Pharmacol. 103, 66–76.
Ellman, G.L., Courtney, K.D., Andres Jr., V., Featherstone, R.M., 1961. A new and rapid colorimetric determination of acetylcholinesterase activity. Biochem. Pharmacol. 7, 88–95.
Foxenberg, R.J., Ellison, C.A., Knaak, J.B., Ma, C., Olson, J.R., 2011. Cytochrome P450-specific human PBPK/PD models for the organophosphorus pesticides: chlorpyrifos and parathion. Toxicology 285, 57–66.
Garabrant, D.H., Aylward, L.L., Betent, S., Chen, Q., Timchalk, C., Burns, C.J., Hays, S.M., Albers, J.W., 2009. Cholinesterase inhibition in chlorpyrifos workers: characterization of biomarkers of exposure and response in relation to urinary TCPγ. J. Expo. Sci. Environ. Epidemiol. 19, 634–642.

Hunter, P.L., Marshal, R.S., Padilla, S., 1997. Automated instrument analysis of cholinesterase activity in tissues from carbamate-treated animals: a cautionary note. Toxicol. Meth. 7, 43–53.

Kousba, A.A., Poet, T.S., Timchalk, C., 2003. Characterization of the *in vitro* kinetic interaction of chlorpyrifos–oxon with rat salivary cholinesterase: a potential biomonitoring matrix. Toxicology 188, 219–232.

Liu, J., Olivier, K., Pope, C., 1999. Comparative neurochemical effects of repeated methyl parathion or chlorpyrifos exposures in neonatal and adult rats. Toxicol. Appl. Pharmacol. 158, 186–196.

Lotti, M., 1995. Cholinesterase inhibition: complexities of interpretation. Clin. Chem. 41 (12 part 2), 1814–1818.

Ma, T., Chambers, J.E., 1994. Kinetic parameters of desulfuration and dearylation of parathion and chlorpyrifos by rat liver microsomes. Food Chem. Toxicol. 32, 763–767.

Marty, M.S., Domoradzki, J.Y., Hansen, S.C., Timchalk, C., Bartels, M.J., Mattsson, J.L., 2007. The effect of route, vehicle, and divided doses on the pharmacokinetics of chlorpyrifos and its metabolite trichlorpyridinol in neonatal Sprague–Dawley rats. Toxicol. Sci. 100, 360–373.

Mattsson, J.L., Johnson, K.A., Albee, R.R., 1986. Lack of neuropathologic consequences of repeated dermal exposure to 2,4-dichlorophenoxyacetic acid in rats. Fund. Appl. Toxicol. 6, 175–181.

Mattsson, J.L., Charles, J.M., Yano, B.L., Cunny, H.C., Wilson, R.D., Bus, J.S., 1997. Single-dose and chronic dietary neurotoxicity screening studies on 2,4-dichlrophenoxyacetic acid in rats. Fundam. Appl. Toxicol. 40, 111–119.

Mattsson, J.L., Maurissen, J.P., Nolan, R.J., Brzak, K.A., 2000. Lack of differential sensitivity to chloinesterase inhibition in fetuses and neonates compared to dams treated perinatally with chlorpyrifos. Toxicol. Sci. 53, 438–446.

Maxwell, D.M., 1992a. The specificity of carboxylesterase protection against the toxicity of organophosphorus compounds. Toxicol. Appl. Pharmacol. 114, 306–312.

Maxwell, D.M., 1992b. Detoxification of organophosphorus compounds by carboxylesterases. In: Chambers, J.E., Levi, P.E. (Eds.), Organophosphates Chemistry, Fate and Effects. Academic Press, New York, pp. 183–199.

Morgan, E.W., Yan, B., Greenway, D., Parkinson, A., 1994. Regulation of two rat liver microsomal carboxylesterase isozymes: species differences, tissue distribution, and the effects of age, sex, and xenobiotic treatment of rats. Arch. Biochem. Biophys. 315, 513–526.

Mortensen, S.R., Chanda, S.M., Hooper, M.J., Padilla, S., 1996. Maturational differences in chlorpyrifos–oxonase activity may contribute to age-related sensitivity to chlorpyrifos. J. Biochem. Toxicol. 11, 279–287.

Moser, V.C., 1995. Comparisons of the acute effects of cholinesterase inhibitors using a neurobehavioral screening battery in rats. Neurotox. Teratol. 17, 617–625.

Moser, V.C., Padilla, S., 1998. Age- and gender-related differences in the time course of behavioral and biochemical effects produced by oral chlorpyrifos in rats. Toxicol. Appl. Pharmacol. 149, 107–119.

Moser, V.C., Chanda, S.M., Mortensen, S.R., Padilla, S., 1998. Age- and gender-related differences in sensitivity to chlorpyrifos in the rat reflect developmental profiles of esterase activities. Toxicol. Sci. 46, 211–222.

Moser, V.C., 2000. Dose–response and time-course of neurobehavioral changes following oral chlorpyrifos in rats of different ages. Neurotoxicol. Teratol. 22, 713–723.

Nostrandt, A.C., Padilla, S., Moser, V.C., 1997. The relationship of oral chlorpyrifos effects on behavior, cholinesterase inhibition, and muscarinic receptor density in rat. Pharmacol. Biochem. Behav. 58, 15–23.

Pope, C.N., Chakraborti, T.K., 1992. Dose-related inhibition of brain and plasma cholinesterase in neonatal and adult rats following sublethal organophosphate exposures. Toxicology 73, 35–43.

Pope, C.N., Chakraborti, T.K., Chapman, M.L., Farrar, J.D., Arthun, D., 1991. Comparison of in vivo cholinesterase inhibition in neonatal and adult rats by three organophosphorothioate insecticides. Toxicology 68, 51–61.

Pope, C.N., Karanth, S., Liu, J., Yan, B., 2005. Comparative carboxylesterase activities in infant and adult liver and their sensitivity to chlorpyrifos oxon. Regul. Toxicol. Pharmacol. 42, 64–69.

Reiss, R., Neal, B., Lamb, J., 2012. Acetylcholinesterase Inhibition Dose–Response Modeling for Chlorpyrifos and Chlorpyrifos–oxon. Regul. Toxicol. Pharmacol. 63, 124–131.

Smith, J.N., Timchalk, C., Bartels, M.J., Poet, T.S., 2011. In vitro age-dependent enzymatic metabolism of chlorpyrifos and chlorpyrifos–oxon in human hepatic microsomes and chlorpyrifos–oxon in plasma. Drug Metab. Dispos. 39, 1353–1362.

Sultatos, L.G., 1994. Mammalian toxicology of organophosphorus pesticides. J. Toxicol. Environ. Health 43, 271–289.

Sultatos, L.G., Shao, M., Murphy, S.D., 1984. The role of hepatic biotransformation in mediating the acute toxicity of the phosphorothionate insecticide chlorpyrifos. Toxicol. Appl. Pharmacol. 73, 60–68.

Szabo, J.R., Young, J.T., Grandjean, M., 1988. Chlorpyrifos: 13-week dietary study in Fischer 344 rats. Summarized in: chlorpyrifos (Dursban®, Lorsban®) dietary exposure assessment. Health Assessment Section, Medical Toxicology Branch. Department of Pesticide Regulation. California Environmental Protection Agency. May 8, 1992.

Timchalk, C., Nolan, R.J., Mendrala, A.L., Dittenber, D.A., Brzak, K.A., Mattsson, J.L., 2002. A physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate insecticide chlorpyrifos in rats and humans. Toxicol. Sci. 66, 34–53.

Timchalk, C., Poet, T.S., Kousba, A.A., 2006. Age-dependent pharmacokinetic and pharmacodynamic response in preweanling rats following oral exposure to the organophosphate insecticide chlorpyrifos. Toxicology 220, 13–25.

US EPA, 2008. United States Environmental Protection Agency. Chlorpyrifos Reassessment. Appendix B – Mode of Action: Inhibition of Acetylcholinesterase (AChE). Health Effects Division, Office of Pesticide Programs, August 27, 2008.

US EPA, 2011. United States Environmental Protection Agency. Chlorpyrifos preliminary human health risk assessment. DP No. D388070. Office of Chemical Safety and Polution Prevention. June 30, 2011. Available from: <http://www.epa.gov/oppsrrd1/registration_review/chlorpyrifos/EPA-HQ-OPP-2008-0850-DRAFT-0024%5B1%5D.pdf> (accessed on 12.02.12).

US EPA Scientific Advisory Panel, 2008. A set of scientific issues being considered by the Environmental Protection Agency regarding: the Agency's evaluation of the toxicity profile of chlorpyrifos. Arlington, VA; September 16–18, 2008, SAP minutes no. 2008-04. Available from: <http://www.epa.gov/hsrb/files/1e2-sap-meeting-minutes-121708.pdf> (accessed on 12.02.2012).

Vidair, C.A., 2004. Age dependence of organophosphate and carbamate neurotoxicity in the postnatal rat: extrapolation to the human. Toxicol. Appl. Pharmacol. 196, 287–302.

Zheng, Q., Olivier, K., Won, Y.K., Pope, C.N., 2000. Comparative cholinergic neurotoxicity of oral chlorpyrifos exposures in preweanling and adult rats. Toxicol. Sci. 55, 124–132.

Message

From: Cindy Smith [csmith@gowanco.com]
Sent: 4/11/2018 1:00:33 PM
To: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
CC: janet collins [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usera98e8fe5]
Subject: Re: Meeting tomorrow

We are held up in security downstairs

> On Apr 11, 2018, at 8:19 AM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:
>
> Yes, you may want to try Kaitlin, Derrick and Venus. Kaitlin and Derrick and I will be in a meeting til 9, but I'm sure someone will be watching the phone.
>
> _____
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> Ex. 6
> beck.nancy@epa.gov
>
> -----Original Message-----
> From: Janet Collins [mailto:jcollins@croplifeamerica.org]
> Sent: Tuesday, April 10, 2018 6:40 PM
> To: Beck, Nancy <Beck.Nancy@epa.gov>
> Cc: csmith@gowanco.com
> Subject: RE: Meeting tomorrow
>
> Thanks very much. See you at 9:00. Do we ask for Kaitlin?
>
> Janet
> Ex. 6
>
> -----Original Message-----
> From: Beck, Nancy [mailto:Beck.Nancy@epa.gov]
> Sent: Tuesday, April 10, 2018 6:03 PM
> To: Janet Collins <jcollins@croplifeamerica.org>
> Cc: csmith@gowanco.com
> Subject: RE: Meeting tomorrow
>
> Of course.
> It will be me, Charlotte Bertrand, Rick Keigwin and Kaitlin Keller.
>
> See you in the morning!
>
> _____
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> Ex. 6
> beck.nancy@epa.gov
>
> -----Original Message-----
> From: Janet Collins [mailto:jcollins@croplifeamerica.org]
> Sent: Tuesday, April 10, 2018 5:16 PM
> To: Beck, Nancy <Beck.Nancy@epa.gov>
> Cc: csmith@gowanco.com
> Subject: Meeting tomorrow
>
> Nancy- can you tell me who will be participating from your shop?
>
> Thanks- see you tomorrow.

Message

From: Michael L. Dourson [Ex. 6]
Sent: 5/16/2018 4:03:16 PM
To: thomas.hornshaw@illinois.gov [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=31c80d85ae22472e97b4e678876dd054-thomas.hornshaw@illinois.gov]; White, Kimberly [Kimberly_White@americanchemistry.com]; sminick@txbiz.org; Pam.Giblin@bakerbotts.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=3820fc6aafd744aca022a05379c52288-Pam.Giblin@bakerbotts.com]; Rick Reiss [rreiss@exponent.com]; roger.brewer@doh.hawaii.gov [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=0751cf70ccd24e67923cf60b935cdbc5-roger.brewer@doh.hawaii.gov]; Anne LeHuray [alehuray@naphthalene.org]; Elizabeth Becker [ebecker@cermonline.com]; Richard Reising [rreising@cermonline.com]; Natalia_Foronda@moh.govt.nz; Gift, Jeff [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=746b029cd80e437d9f62708c339a9ec8-Gift, Jeff]; Ohanian, Edward [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=f119491e2ba8476381a39c57a456ac55-EOhanian]; Rak, Drew [andrew.rak@noblis.org]; Short_Brian [Short_Brian@Allergan.com]; cindy.hafner@epa.state.oh.us [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=ba3a16a2053a4aaf963712c9a51de3c7-cindy.hafner@epa.state.oh.us]; CFChaisson@TheLifeLineGroup.org [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=cda1bb20a2984d49867bd55cef2f8550-CFChaisson@TheLifeLineGroup.org]; Barbara Beck [bbeck@gradientcorp.com]; Kirby Tyndall [kirby.tyndall@pbwllc.com]; svia@awwa.org [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=837e1d66b58a4ea99e240f18e13c4c86-svia@awwa.org]; aoller@nipera.org; Shelley (MPCA) [Shelley.Burman@state.mn.us]; Henry Roman [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=userc094bf11]; Eric Ruder [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usere39539f8]; kathy.hughes@hc-sc.gc.ca [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=3ebcd5751474447aa28a7fbee2b0f882-kathy.hughes@hc-sc.gc.ca]; ted@TedSimon-Toxicology.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=94b4028cf7c342488de5bf9c47e6a304-ted@TedSimon-Toxicology.com]; Lesa Aylward [laylward@summittoxicology.com]; shays@summittoxicology.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=aa0c42dea1244b538f859f31dc5718ba-shays@summittoxicology.com]; Rosalind Schoof [rschoof@Environcorp.com]; Sol Bobst [sol@toxsciadvisors.com]; Bill_Gulledge@americanchemistry.com; scampleman@epri.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=623178f97a72466fa795c644b0042b4e-scampleman@epri.com]; sferenc@cpda.com; SBennett@cspa.org; jack_snyder@styrene.org; tucker.helmes@socma.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=c40c6f3be36a425d8428fd4ce886a86e-tucker.helmes@socma.com]; ansellj@personalcarecouncil.org; mjack@gmaonline.org; fkruszewski@cleaninginstitute.org; clark.carrington@fda.hhs.gov; raj.sharma@gapac.com; Julie.Schroeder@ontario.ca; Emilia [ELonardo@gmaonline.org]; dmccarty@dnr.state.ga.us; Kadry, Abdel-Razak [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=acb325507bf7451b9735813cfde5e417-Kadry, Abdel-Razak]; Cogliano, Vincent [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=51f2736376ac4d32bad2fe7cfef2886b-Cogliano, Vincent]; Hakkinen, Pertti (NIH/NLM) [E] [hakkinenp@mail.nlm.nih.gov]; Tiffany Bredfeldt [Tiffany.Bredfeldt@tceq.texas.gov]; Lynn (LH) [LPottenger@dow.com]; Weis, Christopher (NIH/NIEHS) [E] [christopher.weis@nih.gov]; Amy Rosenstein [Ex. 6] Clark, Becki [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=a906e07f1cd143b9a3c2ddab813b8140-Clark, Becki]; craigbeskid@ehcma.org; jarchbo@toronto.ca; erin.hodge@oahpp.ca; Richard.Beauchamp@dshs.state.tx.us; r.jeffrey.lewis@exxonmobil.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=c38c7ff02df7405b868d54a639586df0-r.jeffrey.lewis@exxonmobil.com]; Bette Meek [bmeek@uottawa.ca]; Greg Paoli [gpaoli@risksciencesint.com]; Alan.Stern@dep.state.nj.us [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=4f82fe5258fc43cf8eaeb92ead4f963f-Alan.Stern@dep.state.nj.us]; James Bus [jbus@exponent.com]; Jarabek, Annie [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=8b1de54d48e1429c8129f6499211dbdb-Jarabek, Annie]; drew.rak@noblis.org; Debra Kaden [dkaden@environcorp.com]; CNorman@PattonBoggs.com

[/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=b8744f41e76d4536bf99465ca01933dc-CNorman@PattonBoggs.com]; Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]; Yamada, Richard (Yujiro) [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=4c34a1e0345e4d26b361b5031430639d-Yamada, Yuj]; Lynne Haber [haberlt@ucmail.uc.edu]; Dr. Andy Maier [maierma@ucmail.uc.edu]; Pat McGinnis [mcginnis@tera.org]; Tricia Underwood Ex. 6 Johnson, Mark S CIV USARMY MEDCOM APHC (US) [mark.s.johnson.civ@mail.mil]; Juberg, Daland (DR) [DRJuberg@dow.com]; Oliver, George (GR) [groliver@dow.com]; Risotto, Steve [Steve_Risotto@americanchemistry.com]; Kelly Houston [aeihq@mindspring.com]; wyf9@cdc.gov; Elizabeth Dittman [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=user5e57ae5b]; Prucha, Christopher [cprucha@wm.com]; Raffaele, Kathleen [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=cc48281bbab34bf5bf3ab1a63780d5ca-Kathleen Raffaele]; Jeff Crum [jcrum@hampmathews.com]; Suzanne_Fitzpatrick [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=3ea0cbe73e0543fa9f5e71188cb177c7-Suzanne_Fit]; Carlin, Danielle (NIH/NIEHS) [E] [danielle.carlin@nih.gov]; Craig.Rowlands@ul.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=2faac2cbab564370ae25dc80951fff70-Craig.Rowla]; Uni Blake [BlakeU@api.org]; Cody Wilson [codywilson@coca-cola.com]; Heidi Holst [hholst@ashland.edu]; Paul Hanlon [paul.hanlon@abbott.com]; steven.hermansky@conagra.com; Rick_Becker@americanchemistry.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=f03aaee7f1014aad916f86c53f886717-Rick_Becker@americanchemistry.com]; Roberts,Stephen M [smroberts@ufl.edu]; janet collins [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usera98e8fe5]; Phelka, Amanda [aphelka@nsf.org]; helen.goeden@state.mn.us [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=b567d1e3f235405783e08b0064579be0-helen.goeden@state.mn.us]

CC: Anita.K.Meyer@usace.army.mil [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=ebb41195f2324548959139e28c25b8a6-Anita.K.Meyer@usace.army.mil]; rperona@neptuneinc.org [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=e3cb7a7889be4e398597d55c36c6e5a4-rperona@neptuneinc.org]; Michael Honeycutt [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usere10b9f5f]; Mumtaz, Moiz (ATSDR/DTHHS/OD) [mgm4@cdc.gov]; Annette Dietz [acdietz@pdx.edu]; Michael Dourson [dourson@tera.org]

Subject: Beyond Science and Decisions Workshop Series

Dear Colleagues

As many of you know, I have rejoined TERA and we are in the process of reestablishing some of its previous collaborative work. Several of you have expressed interest in continuing this workshop series, based on the NAS 2009 document of the same name, and the steering committee of the Alliance for Risk Assessment is favorable.

Part of this restart will be to ask for members of our interested community to serve on a project steering committee. If you or one of your colleagues is interested in serving, please let me know. We would like to have a first conference call of this committee during the last week of May. Our next workshop would be anticipated in the fall of this year. As before, if you have methods on which you might wish some feedback, please feel free to submit them to this committee. With over 40 case studies, this workshop series and its expert scientific panel, might be a good place to showcase your work.

Please feel to share this announcement with interested colleagues. For details associated with this project, see https://tera.org/Alliance%20for%20Risk/ARA_Dose-Response.htm.

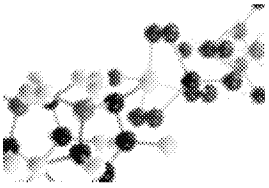
Cheers!

Michael

TERA

Toxicology Excellence For Risk Assessment

—Dedicated to the best use of toxicity data for risk assessment: see http://tera.org/ART/index.html



Current dose response framework information at: https://med.uc.edu/eh/centers/rsc/risk-resources/dose-response-framework

——

Last Workshop

Beyond Science & Decisions - Workshop 9

June 9-10, 2015

Cincinnati, Ohio – University of Cincinnati, Kettering Laboratory Auditorium

Almost 60 organizations have come together under the aegis of the Alliance for Risk Assessment (*ARA*), to conduct a series of meetings, with the ultimate goal of consensus among the participants on a methods compendium highlighting key considerations for applying dose-response techniques for common risk assessment applications. The last workshop (Workshop IX) was held June 9-10 in Cincinnati, Ohio. The workshop was open to the public and broadcast via webinar.

Case Studies that were discussed and sponsors included:

- Adverse outcome pathway (AOP) for Cancer Effects Reported following Oral Exposure to Inorganic Arsenic – ½ day - EPRI
- AOP for Noncancer Effects Reported following Oral Exposure to Inorganic Arsenic – ¼ day - EPRI
- A risk assessment methodology for flame retardants – the FR framework – ½ day – ICL Industrial Products
- AOP for a mutagenic MOA for Cancer (AFB1 and HCC) – ¼ day – ARASP (Center for **A**dvancing **R**isk **A**ssessment **S**cience and **P**olicy) and the US EPA.

Message

From: Jay Vroom [JVroom@croplifeamerica.org]
Sent: 4/13/2018 10:48:53 PM
To: Keigwin, Richard [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=151baabb6a2246a3a312f12a706c0a05-Richard P Keigwin Jr]; Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
Subject: ESA letter from Defenders, CropLife and others
Attachments: ESA FIFRA MOA Letter 041018.pdf

Dear Nancy and Rick,

I wanted to be sure you saw this letter that was finalized and delivered this week. Please let me know if you have any questions.

Best regards,

Jay

Jay Vroom
President & CEO
CropLife America
1156 15th Street, NW
Suite 400
Washington, DC 20005

Ex. 6

Email: vroom@croplifeamerica.org
Executive Assistant: Mary Jo Tomalewski (202.872.3849, mjtomalewski@croplifeamerica.org)

April 10, 2018

The Honorable Ryan Zinke
Secretary
U.S. Department of the Interior
1849 C Street, N.W.
Washington, D.C. 20240
exsec@ios.doi.gov

The Honorable Wilbur Ross
Secretary
U.S. Department of Commerce
1401 Constitution Avenue, N.W.
Washington, D.C. 20230
WLRoss@doc.gov

The Honorable Scott Pruitt
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460
Pruitt.scott@Epa.gov

The Honorable Sonny Perdue
Secretary
U.S. Department of Agriculture
1400 Independence Ave S.W.
Washington, D.C. 20250
Sonny.Purdue@osec.usda.gov

Via Electronic Mail

Re: January 31, 2018 Memorandum of Agreement Implementation

Secretaries Perdue, Ross and Zinke and Administrator Pruitt:

We write to present a unified voice on the opportunity to address one of the most challenging issues facing the intersection of federal pesticide regulation and endangered species conservation: the need for an efficient regulatory process for aligning federal pesticide registration decisions under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) with the requirements of the Endangered Species Act (ESA). We believe these thoughts are both specific and timely as you implement the January 31, 2018 Memorandum of Agreement on Establishment of an Interagency Working Group to Coordinate Endangered Species Act Consultations for Pesticide Registrations and Registration Review (MOA), which we support. For too long, this issue has been marked by divisiveness and conflict as to possible product effects on endangered species and regulatory uncertainty for pesticide manufacturers, farmers, and other users. Your agencies can redouble their efforts from the last four years to move past these conflicts by prioritizing a series of administrative improvements to how pesticides are evaluated. The recent MOA can further this goal considerably.

As a group of diverse stakeholders who care deeply about harmonizing endangered species conservation with agriculture and pest control, we believe that your agencies can and should make further administrative improvements, consistent with the collaborative approaches they have announced, and with their engagement with stakeholders during recent years. There are numerous ways to improve the process of assessing potential impacts to endangered species associated with pesticide registrations. The recommendations here are ones that we mutually support, that we believe are feasible to implement, and that can meaningfully improve the

process. And in pursuing these recommendations, we urge you to engage stakeholders in an open and transparent manner, as contemplated by the MOA.

1. Develop interagency processes on pesticide consultations that enable the EPA, Services, and USDA to make the best use of each agency's expertise and limited resources

The expertise needed to complete robust pesticide consultations already exists within the agencies and should be leveraged to its fullest extent. The U.S. Environmental Protection Agency (EPA) has expertise in ecological risk assessments for pesticides, including risk assessment methods needed to evaluate the potential risks of pesticides to non-target wildlife, such as exposure modeling and probabilistic tools, and requires significant amounts of data for pesticide registrations. The U.S. Fish and Wildlife Service and the National Marine Fisheries Service (collectively, the Services) have substantial expertise on threatened and endangered species, including species biology, distribution, threats, and recovery needs. And the U.S. Department of Agriculture (USDA) has expertise on how pesticides are used in agriculture, including the timing and location of pesticide applications. This use information can be shared with other agencies in ways that do not compromise landowner privacy or specific species locations.

To make better use of limited agency resources, EPA should play a larger role in assessing the potential effects of pesticides on endangered species, including at the population and species levels. For the EPA to play such a role, and other agencies to leverage their existing data and resources, your agencies should start by assessing the effectiveness of existing interagency agreements and guidance on how to complete pesticide consultations. This effort should help ensure that all four agencies have a common understanding of their own responsibilities, the key scientific and policy assumptions that underlie an ESA pesticide consultation, including risk-assessment endpoints, and the data and analyses needed to achieve those endpoints. This assessment would also provide stakeholders with the transparency and accountability that should allow them to support this proposed approach.

New guidance could identify clearer roles for each agency based on expertise and available and reliable data. For example, USDA could be relied on for the cropping and pesticide use data it already collects; EPA for quantitative risk assessment tools and uncertainty analysis; and the Services for defining species ranges and evaluating effects at the species level. At the same time, guidance could also identify ways for the agencies to continue improving collaboration so that one agency is not "handing off" its analysis to another agency, but rather coordinating with that agency throughout the consultation process. An improved approach could also allow stakeholders to provide more information and data during the process, similar to how other endangered species reviews under the ESA are completed.

Your agencies can build additional guidance today and implement it as a living document that can be updated easily to reflect improved methods your agencies develop in the future. If successful, the guidance will help ensure that capable agency scientists—whether sitting at the

EPA or the Services—can share and implement a common understanding of how to perform pesticide consultations, facilitating their collaboration.

2. Use more refined species location maps and better pesticide use data

By using more refined data on where species are likely to occur, the EPA and the Services can improve the occurrence maps of many species compared to some of the maps the Services currently use, many of which are county-level. Refined range maps, which could be produced using species distribution models and other robust scientific approaches, would more accurately depict the true distribution of species and may result in fewer overlaps with areas affected by pesticide use, allowing for a better understanding of potential exposure to those species. This should expedite endangered species review for pesticides, improving the EPA's and the Services' ability to meet statutory timeframes under FIFRA and the ESA.

By further involving pesticide registrants and the public, and considering available data, your agencies can make use of more realistic information on when and how pesticides are applied, thus enabling a more refined assessment. This information, when combined with refined species range maps, may enable the EPA and the Services to identify more instances where pesticide use does not overlap with species habitat. We see promising opportunities to work with USDA, state agencies, species expert organizations, growers, and registrants to improve data on pesticide use patterns.

3. Adopt better endangered species exposure assessments

Better exposure assessments can help the Services and EPA make defensible, science-based conclusions that pesticide exposure is low or absent. One approach is to develop and implement an interagency plan to refine hydrological and other exposure models that adopt more accurate assumptions about endangered species exposure to pesticides. We see opportunities to further refine commonly used models to distinguish between realistic and improbable exposure scenarios. More realistic scenarios would help ensure that conservation efforts focus on the species that are most likely to be affected by potential pesticide exposure.

4. Take advantage of avoidance and minimization opportunities to improve the efficiency and effectiveness of pesticide consultations

EPA's registration of pesticides currently includes requirements to avoid and minimize impacts to non-target organisms. To enhance endangered species review, pesticide registrants could choose to voluntarily adopt additional site-specific avoidance and minimization measures for endangered species as part of EPA's registration process or during consultations. Refined species occurrence data are important to these efforts because they may allow pesticide registrants, farmers, and other users to target protective measures to areas where species and their habitats are likely to occur. They may also result in more pesticide consultations being expeditiously resolved. Such an outcome would represent a win for conservation and for

regulated entities: fewer species potentially exposed to pesticides that could pose a risk to them, and quicker and more predictable pesticide registration decisions.

5. Support opportunities to use voluntary conservation in pesticide evaluations

In addition to avoidance and minimization, a pesticide registrant may choose to consider voluntary conservation efforts as an option to expedite, supplement, or simplify endangered species review for a pesticide. This type of conservation effort (similar to a concept known as compensatory mitigation in other contexts and referred to as "mitigation" below) can also conserve species while expediting or simplifying pesticide consultations. This approach has not played a prominent role in pesticide consultations to date. But if registrants choose to pursue this option, effective and timely conservation efforts consistent with mitigation goals could lead to more efficient consultations in some circumstances.

We urge your agencies to devote resources to help interested stakeholders establish voluntary conservation projects and to integrate those projects into pesticide consultations at the request of registrants. Specifically, we encourage the agencies to work with stakeholders to develop a regulatory framework that further incentivizes voluntary conservation to improve or increase habitat for endangered species.

6. Prioritize species-use combinations for formal consultation

We recommend that your agencies consider developing decision systems to help distinguish among situations that pose low, medium, and high likelihood of jeopardy or adverse modification (JAM) in formal consultation. In developing this system, your agencies could consider both species and pesticide use factors. For example, species factors could include abundance, biological status, and prey base. And use factors could include mode of action, route of entry, and areas of use.

Identifying low, medium, and high-risk scenarios will help your agencies apply the most efficient methods to complete JAM analyses. For many scenarios, proxy measures or general principles of conservation biology and ecotoxicology may be adequate to inform the JAM analysis. For other, higher-risk scenarios, more detailed species- and pesticide-specific analyses may be warranted. The goal should be to complete the JAM analysis for low risk scenarios using efficient yet defensible methods, so that agency staff can focus their limited resources on higher risk scenarios that required more detailed, resource-intensive methods.

We believe that these recommendations for managing endangered species review of pesticides will provide for a more efficient approach to species conservation while providing a sound basis for decisionmaking. We also understand that your agencies would need additional resources and funding to implement the recommendations effectively and expeditiously. We ask for a commitment at the highest levels within your agencies to prioritize these improvements to endangered species review of pesticides. With that commitment, we believe an enduring

solution is possible to the current concerns with the adequacy of endangered species assessments in pesticide consultations.

Sincerely,

CropLife America
Defenders of Wildlife
American Soybean Association
Minor Crop Farmer Alliance
National Association of Corn Growers
National Association of Wheat Growers

cc: Mr. Ray Starling
Special Assistant to the President for Agriculture, Trade and Food Assistance
Raymond.A.Starling@who.eop.gov

Mr. Michael J. Hickey
Chief, Environment Branch, Office of Management and Budget
mhickey@omb.eop.gov

Mr. Chris Prandoni
Associate Director for Natural Resources, Council on Environmental QualityChristopher.D.Prandoni@ceq.eop.gov

Mr. Greg Sheehan
Principal Deputy Director, U.S. Fish and Wildlife Service
Gregory_sheehan@fws.gov

Mr. Chris Oliver
Assistant Administrator for Fisheries, NOAA Fisheries
Chris.W.Oliver@noaa.gov

Ms. Charlotte Bertrand
Acting Principal Deputy Assistant Administrator, EPA Office of Chemical Safety and Pollution Prevention
Bertrand.Charlotte@epa.gov

Dr. Sheryl Kunickis
Director of Office of Pest Management, U.S. Department of Agriculture
Sheryl.Kunickis@osec.usda.gov

Message

From: Janet Collins [jcollins@croplifeamerica.org]
Sent: 4/9/2018 11:08:22 AM
To: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
CC: csmith@gowanco.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=85a93dee627e495997f325593ed303eb-csmith@gowa]; Courtney DeMarco [cdemarco@croplifeamerica.org]
Subject: Update Meeting Materials
Attachments: Chlorpyrifos Epi Review Grasdient 02 15 18.pdf; Debbie Edwards et al 02 18.pdf; Reiss 2015.pdf; Chlorpyrifos_Meas_Error_Comments_215-8248.pdf; Jan 25 2013 Bradbury Letter.pdf

Importance: High
Flag: Flag for follow up

Nancy- again we really appreciate you joining our Strategic Oversight Council discussion on January 25th. You may recall that we had several items we committed to follow up on for you.

1. You raised the concept of a 3rd party review of the epidemiological data that is the basis for EPA's HED Memorandum that reapplies an FQPA 10x to all organophosphate risk assessments. We wanted to highlight for you that some 3rd party reviews of those data have been conducted. I have highlighted the summaries of the following papers and provided them in their entirety if you want to review them:
    a. Debbie Edwards Paper
    b. Rick Reiss/Michael Goodman Paper
    c. Gradient Paper (2015)
2. We pointed out that EPA had completed risk assessments for some organophosphates after the epidemiological data were available to them where no FQPA 10x was applied. Here are the specific examples. Here are the specific examples of organophosphates which EPA removed the FQPA 10x and did not reapply until the HED memo issued in 2015:
    a. Bensulide – EPA scoping document in 2008 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data
    b. Phosmet – EPA scoping document in 2009 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data

    Additionally, I also include a letter written in 2013 by Steve Bradbury (then Director of the Office of Pesticide Programs) regarding the use of these same epidemiology data in risk assessments. It is our contention that for EPA to reapply the FQPA 10x to a compound from which the Agency had removed it, they must have reliable and available data. The researchers have not provided the data to the Agency for the epidemiological studies that are the basis for EPA reapplying the FQPA 10x.

3. Ongoing mechanistic data. You mentioned ongoing research–possibly at ORD –to determine if there is some other mode of action occurring at doses lower than those that inhibit cholinesterase that may cause neurodevelopmental effects. Can you please provide more information about what is being done? We would just like to point that previous discussions on this topic often focused on brain rather than RBC and the focus really needs to be on RBC.

Thanks again for agreeing to meet with us on Wednesday- we appreciate it.

Janet

Ex. 6

Chlorpyrifos and Neurodevelopmental Effects: Overview of the Columbia Study

Decades of research indicates that chlorpyrifos is only toxic at exposures that are high enough to inhibit acetylcholinesterase (AChE) activity in the brain. The range of exposures experienced by children and pregnant women are far lower than those that can cause AChE inhibition.

Even so, several relatively recent epidemiology studies have evaluated prenatal chlorpyrifos exposure and birth outcomes (*e.g.*, infant body weight or head circumference) and neurodevelopmental (*e.g.*, mental and psychomotor) testing results. These studies have found weak and inconsistent associations.

These studies have been critically reviewed several times over the last several years (*e.g.*, Reiss *et al.*, 2015; Edwards *et al.*, 2013; Prueitt *et al.*, 2011; Gradient, 2015). The US EPA Office of Pesticide Programs (OPP) is particularly focused on the Columbia Center for Children's Environmental Health Mothers and Newborn Study (the Columbia study) and has been specifically considering analyses by Dr. Rauh and colleagues published in 2006 (Rauh *et al.*, 2006) and 2011 (Rauh *et al.*, 2011) for its re-evaluation of chlorpyrifos. These studies reported associations between low chlorpyrifos levels *in utero* and lower IQ scores and increased behavioral problems at 3 to 7 years of age. This contradicts the long-standing, strong evidence from toxicology studies demonstrating that these exposure levels do not have neurotoxic effects.

The Columbia study has many strengths compared to other epidemiology studies of chlorpyrifos, but it also has many limitations, many of which have been acknowledged by US EPA and its Scientific Advisory Panel (SAP) (US EPA, 2012). Based on concerns over these limitations, in 2012, the SAP requested additional data and analyses from the Columbia study researchers to evaluate various limitations and then strengthen the reliability of the findings for risk assessment, but the researchers did not comply. Issues related to these studies include:

- These studies relied on only one chlorpyrifos measurement from umbilical cord blood for each child. Using this one measurement, it is not possible to estimate the actual chlorpyrifos exposure experienced during gestation or early childhood.
- Over 40% of children had chlorpyrifos levels that were below the limit of detection (LOD), and over 80% were below the level of validation. To deal with measurements below the LOD, investigators used a statistical approach to estimate the unknown measurements, but this greatly reduced the accuracy of the results.
- 12% of children did not have any cord blood chlorpyrifos measurements, so levels in maternal blood were used as a surrogate measurement. This has similar issues as described in the previous point.
- Children had many other known exposures and lifestyle factors that could have contributed to neurodevelopmental effects that were not accounted for. Although the Columbia investigators attempted to account for some of these factors, it was not possible to fully account for all of them.
- In the 2006 study, the authors stated: "In preliminary analyses, we found no indication of either a linear or nonlinear dose-response relationship between chlorpyrifos levels and developmental outcomes." Associations were only reported when the data were manipulated in a specific way.
- Studies at Mount Sinai Hospital and the University of California (UC) at Berkeley do not confirm the results reported by the Columbia University researchers.

- There is no established biological mode of action to explain the potential neurodevelopmental effects reported. The animal data indicate that dose levels that cause adverse neurodevelopmental outcomes only occur at exposures that inhibit AChE in pregnant rats or offspring. A few subjects in the UC Berkeley study may have had chlorpyrifos concentrations near the lowest estimate for 10% red blood cell (RBC) AChE inhibition. All other study subjects had chlorpyrifos levels that were well below all estimates for RBC AChE inhibition.

We also note that OPP has developed guidelines for evaluating potential measurement error in epidemiology studies for use in risk assessment, but did not adhere to these guidelines for assessing measurement error in epidemiology studies of chlorpyrifos and neurodevelopmental outcomes. Unresolved uncertainties about measurement error in the Columbia study could be addressed with additional analyses of original data, and preliminary quantitative bias analyses of available summary data demonstrate that positive findings in the Columbia study could be explained by exposure or outcome misclassification.

In conclusion, all of the chlorpyrifos epidemiology studies have been reviewed by multiple parties on several occasions over the last decade, including the SAP, and several issues have been brought up repeatedly, but have never been sufficiently addressed by US EPA. Collectively, these studies are not robust enough to change the weight of evidence based on animal toxicity and mechanistic studies.

References

Edwards, D; Juberg, D; Burns, C; Goodman, J; Li, A; Bartels, M; Lickfeldt, D. 2013. "Epidemiology Studies Pertaining to Chlorpyrifos Exposures: Considerations of Reliability and Utility." Report to Dow AgroSciences. Submitted to US EPA, Office of Pesticide Programs. 28p., November 12.

Gradient. 2015. "Measurement Error and Misclassification in Epidemiology Studies of Chlorpyrifos and Neurodevelopmental Outcomes: Submission to US EPA Docket # 2015-05844." Submitted to US EPA Docket. Docket No. 2015-05844. 20p., April 24.

Prueitt, RL; Goodman, JE; Bailey, LA; Rhomberg, LR. 2011. "Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos." *Crit. Rev. Toxicol.* 41(10):822-903. http://informahealthcare.com/doi/pdf/10.3109/10408444.2011.616877.

Rauh, V; Arunajadai, S; Horton, M; Perera, F; Hoepner, L; Barr, DB; Whyatt, R. 2011. "7-Year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide." *Environ. Health Perspect.* 119(8):1196-1201.

Rauh, VA; Garfinkel, R; Perera, FP; Andrews, HF; Hoepner, L; Barr, DB; Whitehead, R; Tang, D; Whyatt, RW. 2006. "Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children." *Pediatrics* 118(6):e1845-e1859.

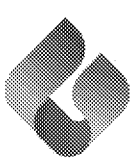
Reiss, R; Chang, ET; Richardson, RJ; Goodman, M. 2015. "A review of epidemiologic studies of low-level exposures to organophosphorus insecticides in non-occupational populations." *Crit. Rev. Toxicol.* 45(7):531-641. doi: 10.3109/10408444.2015.1043976.

US EPA. 2012. "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Chlorpyrifos Health Effects: Minutes of the FIFRA Science Advisory Panel Meeting held on April 10-12, 2012." FIFRA Scientific Advisory Panel, Minutes No. 2012-04, 108p.

Measurement Error and Misclassification in Epidemiology Studies of Chlorpyrifos and Neurodevelopmental Outcomes

Submission to US EPA Docket # 2015-05844

April 24, 2015



GRADIENT

www.gradientcorp.com
20 University Road
Cambridge, MA 02138
617-395-5000

Introduction

In the Revised Human Health Risk Assessment for Chlorpyrifos (HHRA; US EPA, 2014), the United States Environmental Protection Agency (EPA) Office of Pesticide Programs (OPP) reviewed results from several epidemiology studies of prenatal chlorpyrifos exposure and neurodevelopmental outcomes and concluded that chlorpyrifos exposure likely played a role in observed neurodevelopmental effects. Applying principles described in its draft handbook for incorporating epidemiology data in risk assessment, OPP evaluated the potential for measurement error in available epidemiology studies and discussed its potential impacts on measured associations. OPP's evaluation was not sufficiently rigorous to provide a thorough, balanced perspective on the epidemiology literature as a whole. We have identified several specific shortcomings in OPP's assessment of measurement error in individual studies, especially pertaining to analyses of the Columbia study, the cohort to which OPP assigned the greatest weight in the overall evaluation of epidemiology evidence.

After discussing these shortcomings, we describe additional analyses that could be conducted by OPP using raw data from the Columbia study. This additional work could help resolve the remaining uncertainties regarding possible biases caused by measurement error, including residual confounding resulting from measurement error in model covariates. Finally, we present the results of two sensitivity analyses we conducted to estimate some of the potential impacts of measurement error in the Columbia study. These analyses were completed using summary data alone and are, therefore, less informative than the analyses that would be possible if we had access to raw data. Despite this, our results indicate that at least some reported associations could be biased by exposure or outcome misclassification.

Specifically, we discuss the following four points:

1. OPP has developed guidelines for evaluating potential measurement error in epidemiology studies for use in risk assessment;
2. OPP did not adhere to its own guidelines for assessing measurement error in epidemiology studies of chlorpyrifos and neurodevelopmental outcomes;
3. Unresolved uncertainties about measurement error in the Columbia study could be addressed with additional analyses of original data; and
4. Preliminary quantitative bias analyses of available summary data demonstrate that positive findings in the Columbia study could be explained by exposure or outcome misclassification.

1 OPP Has Developed Guidelines for Evaluating Potential Measurement Error in Epidemiology Studies for Use in Risk Assessment

Towards the goal of using the results of epidemiology studies in the "most scientifically robust and transparent way," OPP proposed a draft framework for weighing epidemiology results and integrating them into risk assessment (US EPA, 2010a), and the office solicited comments from the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Science Advisory Panel (SAP) as well as the general public to revise and strengthen the framework.

Several aspects of the framework are relevant to measurements of exposure, outcome, and covariates, and directly or indirectly refer to the impact of measurement error or misclassification on observed

epidemiology associations.[1] Although the framework lacked specific guidance, it indicated that the most useful epidemiology studies for risk assessment employ reliable and valid exposure assessment methods. In addition, OPP stated that one should consider whether covariates relevant to confounding are properly described, measured, and analyzed. Finally, OPP acknowledged that imperfect measurements of exposure or outcome can lead to information bias, which can bias an observed association in either the positive or negative direction. The framework also stated that statistical methods should be evaluated and that epidemiology studies incorporated into risk assessment should include complete descriptions of statistical approaches. This concept applies to measurement error, because choices made in analysis can introduce or amplify biases arising from various types of measurement error, as we demonstrate with specific examples below.

The FIFRA SAP reviewed the framework and commended OPP for developing it (US EPA, 2010b). The SAP reiterated the general importance of establishing clear and robust guidelines for evaluating the quality of epidemiology data, emphasizing "the quality and reliability of the information provided by epidemiologic studies needs to be closely scrutinized" (US EPA, 2010b). To strengthen the framework, SAP recommended that OPP develop a set of specific criteria for determining the acceptability of epidemiology studies, including the use of sensitivity analyses to test the robustness of study results to measurement error in exposure as well as covariates used to adjust for confounding. In a recent publication, EPA scientists from various centers and offices, including OPP, also emphasized the importance of sensitivity analyses when applying epidemiology results to human health risk assessment (Christensen *et al.*, 2015).

In summary, even though the draft framework lacked specific guidance on how measurement error should be assessed and, importantly, how to integrate epidemiology data into risk assessment when measurement error is significant, the framework reinforced EPA's values of transparency and scientific rigor. Feedback from the SAP provided OPP with specific suggestions for incorporating quantitative methods for assessing the impacts of measurement error. Below, we review OPP's evaluation of the epidemiology data used in the chlorpyrifos risk assessment in light of OPP's draft framework and the SAP's feedback on it.

2 OPP Did Not Adhere to Its Own Guidelines for Assessing Measurement Error in Epidemiology Studies of Chlorpyrifos and Neurodevelopmental Outcomes

In the revised HHRA for chlorpyrifos, OPP reviewed 17 peer-reviewed research reports describing prenatal chlorpyrifos exposure and neurodevelopmental outcomes in three children's health cohorts. Of the three cohorts, OPP placed the greatest weight on the results of the "Columbia study," a cohort of minority women and children in New York City, because its exposure assessment was based on cord blood concentrations of chlorpyrifos. By contrast, prenatal exposure in the other two cohorts (the Mt. Sinai study, which also enrolled minority women and children in New York City, and the CHAMACOS study, which included mother-child pairs living in an agricultural community in California) was estimated based on concentrations of chlorpyrifos metabolites in maternal urine. Cord blood chlorpyrifos is considered to be a superior method for estimating prenatal chlorpyrifos exposure for a number of reasons (Prueitt *et al.*, 2011; Eaton *et al.*, 2008). Based largely on the positive associations reported in the Columbia study publications, OPP concluded that chlorpyrifos "likely" played a role in neurodevelopmental outcomes observed in the epidemiology studies.

[1] The terms "measurement error" and "misclassification" refer to errors in continuous and categorical variables, respectively, and are not completely interchangeable. For this report we will use the term "measurement error" for both types, to be concise, except in cases when it is necessary to distinguish between the two types of error.

OPP discussed the general limitations of chlorpyrifos epidemiology studies overall and further detailed the strengths and weaknesses of individual studies in Appendix 3 of the HHRA (US EPA, 2014). Limitations described by OPP included measurement error in estimation of exposures, outcome, and covariates, and, in many cases, OPP qualitatively estimated the potential impacts of error on measured associations. This type of critical evaluation is in line with the OPP framework for incorporating epidemiology results into risk assessment. However, OPP's evaluation overall lacked sufficient rigor and consistency necessary to provide an accurate, unbiased perspective on the results of the epidemiology studies reviewed in the HHRA. Below, we identify critical shortcomings in OPP's assessments of measurement error for each type of measurement (*i.e.*, exposure, outcome, and covariate), and then describe how a more rigorous approach to evaluating these errors would have increased the utility of the epidemiology results in the chlorpyrifos risk assessment.

Exposure Measurement Error

In Section 2.3 of the HHRA (US EPA, 2014), OPP described general factors leading to errors in exposure assessment in all three cohorts. OPP determined that the challenge of estimating accurate chlorpyrifos doses during the most relevant periods of development is a major limitation common to all studies. Sources of measurement error largely stem from the variability in chlorpyrifos biomarker concentrations over short time scales and the fact that exposure assessment in all studies was based on only one biomarker measurement (or, occasionally, two). For these reasons, exposure estimates used in analyses may have differed substantially from true *in utero* exposures experienced during critical developmental periods. In addition, the CHAMACOS and Mt. Sinai studies relied on concentrations of pesticide metabolites (*i.e.*, TCPy and dialkyl phosphates [DAPs]) in maternal urine for exposure assessment. Concentrations of these metabolites have relatively poor specificity for chlorpyrifos exposure, because they reflect exposure to other organophosphate pesticides as well as preformed nontoxic TCPy and DAPs in the environment (Morgan *et al.*, 2005; Lu *et al.*, 2005). Therefore, exposure estimates used in the CHAMACOS and Mt. Sinai studies were affected by additional sources of measurement error beyond those in the Columbia study.

The HHRA further expanded on aspects of measurement error in its detailed reviews of individual studies (US EPA, 2014, Appendix 3), and, in all cases, it predicted that exposure measurement error was nondifferential with respect to neurodevelopmental outcomes. OPP repeatedly stated that because the errors are nondifferential, they likely biased observed epidemiology associations towards the null, thereby masking any true relationships. While it is true that nondifferential exposure measurement error will often have this effect on measured associations, it is not always the case. Nondifferential error is guaranteed to bias associations towards the null only under specific conditions, none of which were critically assessed by OPP or individual chlorpyrifos epidemiology researchers.

Several quantitative analyses have demonstrated realistic scenarios under which approximately nondifferential exposure measurement errors can bias results away from the null (Flegal *et al.*, 1991; Jurek *et al.*, 2008; Dosemeci *et al.*, 1990). Jurek *et al.* (2008) showed that associations measured in datasets with low exposure prevalence are especially vulnerable to exposure misclassification that is nearly, but not completely, nondifferential. Overall, the possibility that exposure measurement error can bias results away from the null should not be dismissed by OPP, especially considering that it is common practice for researchers to run multiple statistical models and selectively present the results of models that yield positive findings. In fact, model selection bias is evident in the Columbia study, as discussed below.

OPP also failed to consider that the practice of categorizing continuous exposure measurements can lead to exposure misclassification. Statisticians generally discourage using categories when continuous values

are available, because doing so results in lower variability and, subsequently, reduced statistical power to detect true effects (Froslie *et al.*, 2010). Another disadvantage of this approach is that categorizing continuous exposure measurements affected by nondifferential errors can lead to differential misclassification errors, potentially biasing observed associations in either direction (Flegal *et al.*, 1991). In all three children's health cohorts, continuous exposure measurements were grouped into categories for at least some analyses (Barr *et al.*, 2010; Eskenazi *et al.*, 2004, 2007; Engel *et al.*, 2007; Berkowitz *et al.*, 2004; Rauh *et al.*, 2006; Lovasi *et al.*, 2011), but OPP did not mention any potential biases associated with this approach.

In addition, there can be more serious consequences of categorizing continuous measures of exposure. A close look at the statistical methods employed in the Columbia study indicate that categorization of continuous chlorpyrifos measurements most certainly biased study findings away from the null. Rauh *et al.* (2006) estimated adjusted odds ratios (ORs) of 2.37 (95% confidence interval [CI]: 1.08-5.19) and 4.52 (95% CI: 1.61-12.70) for mental and psychomotor delay, respectively, in association with high *versus* low chlorpyrifos exposure. However, their method of defining high and low exposure groups likely contributed to false positive results. In the Methods section, the authors stated that preliminary analyses of Mental Development Index (MDI) and Psychomotor Development Index (PDI) scores indicated neither a "linear or nonlinear dose-response relationship between chlorpyrifos levels and developmental outcomes," but they provided no details about how this was determined (Rauh *et al.*, 2006).

Next, they explored associations across various categories of exposure. Continuous chlorpyrifos levels were categorized into four groups, consisting of concentrations that were less than the limit of detection (LOD) (n = 80) and tertiles of those that were detectable (*i.e.*, first tertile, n = 65; second tertile, n = 39; and third tertile, n = 44). Rauh *et al.* (2006) calculated effect estimates for each category and observed that the strongest associations resulted when exposure groups were redefined in a dichotomous manner, with low and high exposure groups defined as below and above 6.17 pg/g, respectively, *i.e.*, the concentration cut-off between the third and fourth highest categories of exposure. This description of preliminary results appeared only in the Methods section of the article, as an explanation for the choice of the 6.17 pg/g cut-point to define low *vs.* high exposure. By contrast, in the Results section of the article, Rauh *et al.* (2006) mentioned neither the null findings of their preliminary analysis nor the weaker associations observed for alternative categorization schemes.

Another source of exposure measurement error that received little attention in the HHRA was the treatment of nondetectable biomarker readings. Chlorpyrifos cord blood measurements for 43% of the Columbia cohort fell below the LOD. In Rauh *et al.* (2011), the values below the LOD were imputed so that exposure could be analyzed as a continuous variable. Based solely on an assumption of a lognormal distribution, missing values were assigned an expected value based on the distributional shape. That is, the nondetectable chlorpyrifos concentrations were assigned the most likely value expected based on the predicted shape of a log-normal distribution extending below the LOD. In statistics, this expected value is referred to as "$E(X|X < LOD)$."

The advantage of imputing nondetectable chlorpyrifos concentrations is that all subjects can be included in the analyses, and the method of imputation used by Rauh *et al.* (2001) yields unbiased regression results when the proportion of nondetectable values is small. For example, Lubin *et al.* (2004) conducted simulations to critically evaluate various methods for imputing nondetectable exposure measurements and the impacts on subsequent regression analyses. When the proportion of nondetectable values was modest (*i.e.*, 5-10%), substitution of $E(X|X < LOD)$ produced valid results. However, in datasets with larger proportions of nondetectable values, substitution of $E(X|X < LOD)$ resulted in biased coefficients and standard errors, and Lubin *et al.* (2004) concluded that more sophisticated treatments of nondetectable values, such as multiple imputation, is warranted in this scenario.

Rauh *et al.* (2011) gave little attention to the possibility that their treatment of nondetectable chlorpyrifos concentrations may have led to bias. They described a sensitivity analysis in which the regression analysis was repeated with only the detectable chlorpyrifos levels and stated that they observed "no consistent differences in estimates." No data were shown to support this statement, so it is difficult to judge the validity of this argument; also, this sensitivity analysis does not address potential biases in standard errors, which can affect inferences based on regression results. Furthermore, the investigators' conclusion about a lack of a threshold at low doses is undermined by the fact that the lower 43% of the chlorpyrifos concentrations were nondetectable and imputed using a potentially biased approach.

Outcome Measurement Error

In the HHRA, OPP discussed the challenges inherent in accurately measuring neurobehavioral outcomes in young children and acknowledged potential biases that could have affected the results of the studies they reviewed. A single clinical neurobehavioral test result in an individual is not sufficient to consider an outcome as a functional impairment or illness, even if the result is "abnormal." Also, studies assessing single measurements of neurobehavioral outcome at each time point may be subject to additional measurement error as a result of within-subject variability in results (Eaton *et al.*, 2008). There are many extraneous factors that affect the results of clinical tests, such as test administrator training and blinding to exposure status, and the child's physical activity level, diet, medication use, co-exposures, and whether or not they are obese (Eaton *et al.*, 2008). In particular, many of these tests require an advanced level of training and expertise in test administration (Leonard *et al.*, 2001). Taken together, these factors indicate that outcome measurements may have been highly influenced by errors and misclassification.

Despite several limitations in assessment methods, both OPP and the SAP concluded that the chlorpyrifos epidemiology studies they reviewed utilized the "best available" measurement tools and conducted testing in consistent and standardized ways. OPP predicted that most measurement errors in outcome assessment were nondifferential and that, as a result, any bias in the measured epidemiology results likely would have been towards the null. Our previous discussion regarding nondifferential measurement error applies by analogy to outcome measurement as well. Specifically, nondifferential outcome error does not guarantee that bias is towards the null, except under very specific conditions. If outcome error is nearly, but not perfectly, nondifferential in nature, the error or misclassification can bias associations in either direction, and studies with rare outcomes are especially susceptible to this (Jurek *et al.*, 2008). In the case of the three binary outcomes related to behavioral disorders at 36 months analyzed in Rauh *et al.* (2006), a very small number of children were diagnosed as having a behavioral problem, and any outcome misclassification could have biased associations substantially. We explore this possibility in a quantitative bias analysis in Section 4, below.

In addition, several continuous measures of neurodevelopmental outcomes were dichotomized for use in logistic regression, and the choice of cut-points for diagnosing delayed *versus* non-delayed children may have strongly influenced results. In the Columbia study, scores of 85 on the PDI and the MDI were used to distinguish between children who were "normal" *versus* "delayed," but no rationale or citation for this specific cut-point was provided. In contrast, other sources indicate that the typical cut-offs for moderate and severe development delay using the BSID-II are 70 and 55, respectively (Bos, 2013). In the absence of justification for a cut-off of 85, it is plausible that the categories have been defined to maximize positive findings, as was the case for exposure categories, described above.

Finally, OPP briefly noted that differential errors are possible in one outcome assessment tool, the Child Behavior Checklist (CBCL). Aside from a brief mention of this possibility, however, OPP did not discuss the potential impact that utilizing this tool may have had on study findings (Eskenazi *et al.*, 2006, 2007;

Rauh *et al.*, 2006). The CBCL is a survey completed by mothers and is based on subjective judgment of child behavior. It is feasible that mothers would be more or less likely to report behavioral issues based on study results at earlier time points in the follow-up period. For example, mothers who tested high for chlorpyrifos or other chemicals at the initiation of the study may have been more likely to suspect that this exposure could be the cause of behavior disorders and, therefore, may have been more likely to over-report problematic symptoms their child developed at later time points.

If mothers of children in higher exposure categories differentially over-reported child symptoms at age 36 months on the CBCL, effect estimates would have been biased high. Given the very small number of children identified as having behavioral problems based on the CBCL in Rauh *et al.* (2006), the impact of only one or two misclassified children could have had a profound impact on the measured associations. Rauh *et al.* (2006) did not present counts of children with and without behavioral problems in each exposure group, but Table 3 in their paper shows that only 3.4% of the cohort was diagnosed with attention problems at 36 months. This indicates that there were seven children diagnosed with attention problems, out of the cohort of 228 children. This small number of children in the clinical range for attention problems is reflected in the very wide CIs calculated in the multivariate logistic regression (OR = 11.26, 95% CI: 1.79-70.99 for attention problems) and indicates the risk estimate is not stable.

Covariate Measurement Error

The HHRA noted that several important confounding factors may have biased the results of the chlorpyrifos epidemiology studies OPP reviewed in either positive or negative directions. The potential for confounding in these studies is especially high because several maternal characteristics are strongly associated with both exposure and outcome. For example, as acknowledged by OPP in the HHRA, ample research has established that early life neurodevelopment is positively associated with indicators of increased socioeconomic status (SES). Quantitative analyses have demonstrated that epidemiology studies of low dose environmental exposures and neurodevelopmental outcomes can be confounded by maternal intelligence, home environment, and SES, even if differences in these factors between exposure groups are small (Mink *et al.*, 2004). Mink *et al.* (2004) found that substantial confounding can occur even when these variables are measured and included as adjustment variables, because measures of SES are often inaccurate and residual confounding may persist in multivariate regression.

OPP claims that the issue of confounding was addressed, in part, by the restriction of study cohorts to relatively homogeneous populations. However, women enrolled in each of the three cohort studies displayed a substantial amount of heterogeneity in several ways. For example, significant associations between outcomes and multiple maternal characteristics, including environmental tobacco smoke (ETS) exposure, material hardship, and maternal IQ were observed within the Columbia study cohort, demonstrating that confounding was likely (Rauh *et al.*, 2004, 2011). Likewise, in the CHAMACOS cohort, higher DAPs were measured in mothers with lower intelligence and lower HOME scores (*i.e.*, lower measures of the quality of care-taking environment) (Bouchard *et al.*, 2011).

Even though all the chlorpyrifos epidemiology studies employed multivariate analyses to mitigate confounding, measurement error in covariates limits the effectiveness of the statistical adjustments. Maternal smoking, drug use, and drinking during pregnancy were ascertained by self-report, and these stigmatized behaviors are likely under-reported by many women. It is striking that, despite this, the prevalence of drinking during pregnancy in the Columbia cohort was estimated to be 25% (Whyatt *et al.*, 2004), but none of the Columbia study analyses considered confounding by alcohol use. In fact, this maternal behavior was not mentioned in any reports that followed Whyatt *et al.* (2004). OPP indirectly addressed this topic by noting that a small number of women used alcohol in the Columbia cohort and citing the low percentage of women who reported engaging in "heavy drinking." Because a substantial

proportion of the Columbia cohort reported drinking, and prenatal alcohol exposure may confound the relationship between chlorpyrifos and outcomes, reported associations may have been biased.

In addition, data on some covariates were collected as continuous measures but then dichotomized for use in multivariate analysis, thereby artificially reducing the variability and effectiveness of statistical adjustment (Rauh *et al.*, 2006; 2011; Eskenazi *et al.*, 2007). This practice is commonly discouraged by statisticians (Altman, 2006), and this is an important limitation of the Columbia study in particular.

Covariate measurement error is also likely to have limited researchers' ability to determine what factors confound epidemiology associations. Columbia study researchers asserted that certain factors could not confound relationships on the basis of statistical significance testing. For example, they concluded that because correlations between blood lead levels and both exposure and outcome were not significant, lead was not a confounder; OPP agreed with this assessment. However, correlation was assessed in a subsample of only 89 mother-child pairs, and the test was likely underpowered to detect true associations. Neither OPP nor the researchers considered that small sample sizes and measurement error in covariates limited the statistical power to detect true associations. OPP discussed limited samples sizes and measurement error elsewhere in the HHRA, but only as factors that may have masked true associations. In contrast, the HHRA generally did not give attention to the methodological limitations that may have had the opposite effect.

This use of statistical significance testing for assessing whether a certain factor acts as a confounder and, likewise, for selecting covariates to include in adjusted models is generally discouraged by epidemiologists (Rothman *et al.*, 2008). Interestingly, OPP noted this several times in the HHRA in reviews of individual studies, but only applied the criticism to the CHAMACOS and Mt. Sinai studies. Columbia study analyses should be held to the same standard; OPP should more closely scrutinize the decisions researchers made regarding model covariates and the resulting potential for residual confounding.

Shortcomings in OPP's Overall Summary of Measurement Errors in Chlorpyrifos Epidemiology Studies

In the HHRA's overall integration of epidemiology evidence, the relative impacts of exposure measurement error and residual confounding were directly compared. OPP concluded that, even though it is possible that residual confounding biased observed epidemiology associations away from the null, this bias was likely weaker than the effects of nondifferential exposure measurement error, which OPP maintained biased results towards the null. OPP indicated that this argument was supported by a review of exposure measurement error and confounding in occupational epidemiology studies (Blair *et al.*, 2007). However, Blair *et al.* (2007) is not directly applicable to chlorpyrifos epidemiology research, because exposure assessment in occupational settings typically involves record reviews or retrospective self-reports; these methods are much more susceptible to measurement errors, including differential errors such as recall bias. In contrast, the chlorpyrifos epidemiology studies that OPP reviewed in the risk assessment depended on objective biomarker measurements of exposure, which are far less susceptible to differential measurement errors.

More careful consideration of the magnitude and implications of exposure, outcome, and covariate measurement errors in these studies is needed. As we discussed in depth above, there are many ways by which measurement error, even when nondifferential, may have contributed to false positive findings between chlorpyrifos and neurodevelopmental outcomes. Without additional evaluation, OPP's conclusion that chlorpyrifos exposure likely played a role in neurodevelopmental effects is not well supported.

3 Unresolved Uncertainties About Measurement Error in the Columbia Study Could Be Addressed with Additional Analyses of Original Data

Following two SAP reviews of the draft chlorpyrifos HHRA, OPP determined that certain areas of uncertainty limited the incorporation of Columbia study results into the risk assessment (US EPA, 2008, 2012). OPP therefore requested that Columbia study researchers provide the original analytic file used to conduct the analyses reported in Rauh *et al.* (2006; 2011) and Whyatt *et al.* (2004), so that uncertainties could be carefully evaluated. Columbia researchers refused the request, but agreed to meet with OPP researchers to address OPP's concerns. Based on the discussion at this meeting and additional information subsequently provided to OPP on request, OPP dropped its previous request for original data (US EPA, 2014, Appendix 6, p. 384).

However, several key uncertainties were inadequately addressed by the additional information and analyses, and OPP should renew the request to access the studies' original raw data. In this section, we describe several analyses that should be conducted using the raw data to address the remaining uncertainties about measurement error in the Columbia study. Doing so would help OPP achieve its goal of transparency while also critically evaluating epidemiology data utilized in the risk assessment. This is especially important for the Columbia study, given the greater weight it received in the HHRA evaluation.

None of the analyses we suggest below require additional data collection and most are simple to perform, with results that are easy to interpret. Our final suggestion is a more in-depth analysis aimed at accounting for multiple errors and uncertainties simultaneously. This type of quantitative bias assessment would be more difficult to conduct and interpret, but the results would provide critical insights into the potential impact of errors on the chlorpyrifos epidemiology studies that OPP evaluated. Researchers in academia, government, and industry have called for an increased use of such methods to improve the utility of epidemiology data in human health risk assessment (Burns *et al.*, 2014).

Methods of Adjusting for Potential Confounders

OPP maintained that Columbia study researchers addressed the potential for confounding "to the extent possible," but we believe that additional analyses could provide meaningful insight into the magnitude of residual confounding in reported associations. Simple analyses should be conducted to test the assertion of both OPP and Columbia investigators that lead, polycyclic aromatic hydrocarbon (PAH), alcohol, and ETS exposure did not confound the positive associations between chlorpyrifos and neurodevelopment. Rather than rely on correlational analyses and the results of statistical significance testing, the main findings of the Columbia study should be re-analyzed to determine whether results are sensitive to the inclusion of lead levels, PAH exposures, and reported maternal drinking. Lead exposure data were available for only a subset of children, but missing values could be imputed fairly easily using other available characteristics of mothers and children.

Similarly, additional analyses should explore whether factors that have been dichotomized for multivariate analysis (*e.g.*, years of maternal education and household income) have a stronger impact on measured associations when included in the model as continuous variables.

Treatment of Missing Data

OPP should evaluate whether the method used to impute missing chlorpyrifos measurements in Rauh *et al.* (2011) could have biased regression coefficients or affected the size of standard errors. With the original dataset, OPP could assess whether the results of regression are sensitive to variations on the imputation method employed. Simulations conducted by Lubin *et al.* (2004) showed that the regression results were highly sensitive to the methods used to impute exposures below the LOD, especially when the proportion of missing data is relatively large, as is the case in the Columbia study.

Ad hoc Cut-points Chosen for Categorization of Exposure, Outcome, and Covariate Variables

As described above, the Columbia researchers used inappropriate methods to define high- and low-exposure groups for analyses of several neurobehavioral outcomes (Rauh *et al.*, 2006). Similarly, cut-points for the dichotomization of outcomes from continuous measurements appeared to be arbitrary. In fact, it is possible that these two potential sources of bias were compounded in the analyses of dichotomized exposures and outcomes. The small numbers of cases in high-exposure categories (see Table 1) increases the likelihood that minor variations in the cut-points could have substantial impacts on the counts of exposed cases and noncases and, subsequently, on measured associations (Jurek *et al.*, 2008).

To rigorously assess whether key epidemiology findings from the Columbia study are sensitive to cut-points, further analyses should be conducted in which main effects are recalculated for a variety of exposure and outcome category cut-points. An analogous set of sensitivity analyses focused on the cut-points applied to dichotomized confounders, such as years of maternal education, should be conducted as well.

Variations in Subject Characteristics Before and After the Chlorpyrifos Ban

In the Columbia study, exposures to chlorpyrifos dramatically decreased across the 6-year period during which enrolled mothers gave birth. As a result, chlorpyrifos exposure is strongly correlated with calendar time in this cohort. If characteristics of the enrolled subjects varied over time, a false association between chlorpyrifos and neurobehavioral outcomes could have occurred. If recruiting strategies or locations changed across the 6 years of subject enrollment, for example, it is plausible that women recruited later in the study period were consistently higher or lower in SES or some other factor strongly associated with child neurodevelopment. Even though researchers attempted to control for SES-related factors in their analyses, residual confounding was likely to have occurred, for the reasons discussed above. Close inspection of maternal characteristics and patterns over time may indicate that the characteristics of enrolled women shifted over the recruitment period, and this should prompt increased scrutiny on the methods used to control for confounding.

Probabilistic Bias Assessment of Multiple Measurement Errors and Biases

Finally, a detailed quantitative assessment of potential biases should be conducted with the original data from the Columbia study. As an alternative to the deterministic approaches to assessing potential biases one-by-one, described above, probabilistic methods could be employed to explore the impact of exposure and outcome measurement error, selection bias, and unmeasured confounding simultaneously. Use of the original datasets would ensure that the correlation structure is preserved. For this analysis, estimates of the magnitude of differential and nondifferential errors in exposure and outcome measures are assumed,

and uncertainty in these parameters is modeled using distributions of plausible values. Then, Monte Carlo sampling of parameters from probability distributions and reanalysis of the dataset over thousands of iterations produces distributions of effect estimates that reflect uncertainty, bias, and variability in measured associations. Several examples of this approach can be found in the literature, and researchers have called for increased utilization of these methods in epidemiology studies (Maldonado, 2008; Lash and Fink, 2003; Meliker *et al.*, 2010; Lash, 2009).

4 Preliminary Quantitative Bias Analyses of Available Summary Data Demonstrate That False Positive Findings in the Columbia Study Could Be Explained by Exposure or Outcome Misclassification

In the absence of original individual-level data, quantitative bias analyses can be conducted using summary data of the type generally presented in research publications (Lash, 2009). Even though this type of sensitivity analysis is less informative than that which is possible with individual-level data, these analyses can provide important insights into the potential ramifications of measurement error, selection bias, and/or unmeasured confounding on reported associations. We provide two examples of deterministic sensitivity analyses below, both of which were conducted using the summary results provided in Rauh *et al.* (2006). The first is a deterministic analysis illustrating how the choice of exposure cut-points is highly influential on estimated ORs for psychomotor and mental delay associated with various categories of chlorpyrifos exposure. As we discussed above, this variability in ORs may have led to false positive findings if researchers chose cut-points to maximize associations. The second analysis is a deterministic analysis of the potential impact of differential outcome misclassification on reported relationships with CBCL-derived outcomes in Rauh *et al.* (2006).

Sensitivity of Epidemiology Results to Variations in Exposure Categorization

Rauh *et al.* (2006) reported an adjusted OR of 2.37 (95% CI: 1.08-5.19) for mental delay and an adjusted OR of 4.52 (95% CI: 1.61-12.70) for psychomotor delay in association with high *versus* low chlorpyrifos exposure. In response to a request from the FIFRA SAP, the authors provided a more detailed breakdown of cases and noncases across four exposure categories (Table 1).

Using these data, we have calculated crude ORs to explore the variation in ORs observed using four separate schemes of exposure categorization.[2] Calculation of adjusted OR is not possible without raw data, so we have conducted this analysis with unadjusted risk estimates; we expect that similar patterns would result in multivariate analyses as well. As shown in Table 2 and 3, the magnitude and precision of results is highly sensitive to methods used to model exposure. The interpretation of the main findings of Rauh *et al.* (2006) is impacted by this inconsistency in risk estimates across categorization schemes as well as the investigators' selective presentation of the strongest and most precise risk estimates in their publication.

Analysis of Bias from Outcome Misclassification

We conducted a quantitative bias assessment to determine whether risk estimates for the three CBCL-derived outcomes (attention problems,[3] attention deficit hyperactivity disorder [ADHD] problems, and

[2] ORs were calculated in the traditional method using a 2 x 2 table: OR= (# diseased in exposed group)*(# nondiseased in unexposed group)/[(# diseased in unexposed group)*(# diseased in unexposed group)].

[3] Rauh *et al.* (2006) described outcomes as attention, ADHD and PDD "problems" based on the 98th percentile of CBCL scores in each domain, and we use the same wording here.

pervasive developmental disorder [PDD]) reported by Rauh *et al.* (2006) could be affected by outcome misclassification. As we discussed above, a small number of children were classified as having each problem at 36 months of age, ranging from 7 children identified as having attention problems to 11 with PDD problems. Because CBCL results are based on subjective reports of mothers and because parents of children found to be highly exposed to chlorpyrifos during pregnancy may be more likely to report health problems, it is feasible that some exposed children were misclassified as having a behavioral problem.

To explore the impact of misclassification of one, two, or three exposed children on risk estimates, we reconstructed 2 x 2 tables for CBCL outcomes based on percentages presented in Table 3 of Rauh *et al.* (2006). Only adjusted ORs were presented in the paper, but we did not have access to individual-level data and, so, conducted our analysis instead using crude ORs in the same method described above. We expect that impacts of misclassification would be similar in an adjusted analysis.

As shown in Table 4, misclassification of a small number of children in the exposed category would have a dramatic impact on the magnitude of risk estimates for these three outcomes. This indicates that the risk estimates reported by Rauh *et al.* (2006) may be vulnerable to outcome misclassification, even if only 1 or 2 children out of the 228 in the cohort were falsely identified as having a behavioral problem based on subjective parental reports. It is important to note that misclassification of outcome may have been differential, or it may have been nondifferential, but by random chance affected the high exposure category specifically. Regardless of the mechanism, the impact of a very small amount of misclassification would be substantial.

5 Conclusion

OPP established a framework for incorporating epidemiology research into risk assessment, which includes an evaluation of the accuracy of exposure, outcome, and covariate measurements. In the chlorpyrifos HHRA, OPP critically assessed the nature and magnitude of errors in these measurements for individual epidemiology studies, as well as for the body of research as a whole. However, the HHRA's evaluation of the impact of these errors was not sufficiently rigorous or consistent. OPP could greatly improve the utility of the Columbia study epidemiology results for risk assessment by conducting quantitative bias analyses such as those we described in Section 4 or by renewing the request for original data from investigators to pursue more substantial bias analyses, as we described in Section 3. As other researchers have noted, it is crucial to conduct quantitative sensitivity analyses when important policy decisions are to be based on the results of epidemiology research (Jurek *et al.*, 2008; Christensen *et al.*, 2015; Burns *et al.*, 2014).

Acknowledgment

This work was funded by Dow AgroSciences, LLC. This work represents the individual professional views of the authors and not necessarily the views of Dow AgroSciences, LLC.

References

Altman, DG; Royston, P. 2006. "The cost of dichotomising continuous variables." *BMJ* 332(7549):1080. doi: 10.1136/bmj.332.7549.1080.

Barr, DB; Ananth, CV; Yan, X; Lashley, S; Smulian, JC; Ledoux, TA; Hore, P; Robson, MG. 2010. "Pesticide concentrations in maternal and umbilical cord sera and their relation to birth outcomes in a population of pregnant women and newborns in New Jersey." *Sci. Total Environ.* 408(4):790-795.

Berkowitz, GS; Wetmur, JG; Birman-Deych, E; Obel, J; Lapinski, RH; Godbold, JH; Holzman, IR; Wolff, MS. 2004. "*In utero* pesticide exposure, maternal paraoxonase activity, and head circumference." *Environ. Health Perspect.* 112(3):388-391.

Blair, A; Stewart, P; Lubin, JH; Forastiere, F. 2007. "Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures."*Am. J. Ind. Med.* 50(3):199-207.

Bos, AF. 2013. "Bayley-II or Bayley-III: What do the scores tell us (Commentary)?" *Dev. Med. Child Neurol.* 55(11):978-979. doi: 10.1111/dmcn.12234.

Bouchard, MF; Chevrier, J; Harley, KG; Kogut, K; Vedar, M; Calderon, N; Trujillo, C; Johnson, C; Bradman, A; Barr, DB; Eskenazi, B. 2011. "Prenatal exposure to organophosphate pesticides and IQ in 7-year old children." *Environ. Health Perspect.* 119(8):1189-1195.

Burns, CJ; Wright, JM; Pierson, JB; Bateson, TF; Burstyn, I; Goldstein, DA; Klaunig, JE; Luben, TJ; Mihlan, G; Ritter, L; Schnatter, AR; Symons, JM; Yi, KD. 2014. "Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments." *Environ. Health Perspect.* 122(11):1160-1165. doi: 10.1289/ehp.1308062.

Christensen, K; Christensen, CH; Wright, JM; Galizia, A; Glenn, BS; Scott, CS; Mall, JK; Bateson, TF; Murphy, PA; Cooper, GS. 2014. "The use of epidemiology in risk assessment: Challenges and opportunities." *Hum. Ecol. Risk Assess.* doi: 10.1080/10807039.2014.967039.

Dosemeci, M; Wacholder, S; Lubin, JH. 1990. "Does nondifferential misclassification of exposure always bias a true effect toward the null value?" *Am. J. Epidemiol.* 132(4):746-748.

Eaton DL; Daroff RB; Autrup H; Bridges J; Buffler P; Costa LG; Coyle J; McKhann G; Mobley WC; Nadel L; Neubert D; Schulte-Hermann R; Spencer PS. 2008. "Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment." *Crit. Rev. Toxicol.* 38(Suppl. 2):1-125.

Engel, SM; Berkowitz, GS; Barr, DB; Teitelbaum, SL; Siskind, J; Meisel, SJ; Wetmur, JG; Wolff, MS. 2007. "Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton Neonatal Behavioral Assessment Scale in a multiethnic pregnancy cohort." *Am. J. Epidemiol.* 165(12):1397-404.

Eskenazi, B; Marks, AR; Bradman, A; Harley, K; Barr, DB; Johnson, C; Morga, N; Jewell, NP. 2007. "Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children." *Environ. Health Perspect.* 115(5):792-798.

Eskenazi, B; Harley, K; Bradman, A; Weltzien, E; Jewell, NP; Barr, DB; Furlong, CE; Holland, NT. 2004. "Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population." *Environ. Health Perspect.* 112(10):1116-1124.

Eskenazi, B; Marks, AR; Bradman, A; Fenster, L; Johnson, C; Barr, DB; Jewell, NP. 2006. "*In utero* exposure to dichlorodiphenyltrichloroethane (DDT) and dichlorodiphenyldichloroethylene (DDE) and neurodevelopment among young Mexican American children." *Pediatrics* 118(1):233-241.

Flegal, KM; Keyl, PM; Nieto, FJ. 1991. "Differential misclassification arising from nondifferential errors in exposure measurement." *Am. J. Epidemiol.* 134(10):1233-1244.

Froslie, KF; Roislien, J; Laake, P; Henriksen, T; Qvigstad, E; Veierod, MB. 2010. "Categorisation of continuous exposure variables revisited. A response to the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study." *BMC Med. Res. Methodol.* 10:103. doi: 10.1186/1471-2288-10-103.

Jurek, AM; Greenland, S; Maldonado, G. 2008.\ "How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null?" *Int. J. Epidemiol.* 37(2):382-385.

Lash, TL; Fink, AK. 2003. "Semi-automated sensitivity analysis to assess systematic errors in observational data." *Epidemiology* 14(4):451-458. doi: 10.1097/01.EDE.0000071419.41011.cf.

Lash, TL; Fox, MP; Fink, AK. 2009. *Applying Quantitative Bias Analysis to Epidemiologic Data.* Springer, New York, NY. 192p.

Leonard, CH; Piecuch, RE; Cooper, BA. 2001. "Use of the Bayley Infant Neurodevelopmental Screener with low birth weight infants." *J. Pediatr. Psychol.* 26(1):33-40. doi: 10.1093/jpepsy/26.1.33.

Lovasi, GS; Quinn, JW; Rauh, VA; Perera, FP; Andrews, HF; Garfinkel, R; Hoepner, L; Whyatt, R; Rundle, A. 2011. "Chlorpyrifos exposure and urban residential environment characteristics as determinants of early childhood neurodevelopment." *Am. J. Public Health* 101(1):63-70. doi: 10.2105/AJPH.2009.168419.

Lu, C; Bravo, R; Caltabiano, LM; Irish, RM; Weerasekera, G; Barr, DB. 2005. "The presence of dialkylphosphates in fresh fruit juices: Implication for organophosphorus pesticide exposure and risk assessments." *J. Toxicol. Environ. Health A* 68(3):209-227. doi: 10.1080/15287390590890554.

Lubin, JH; Colt, JS; Camann, D; Davis, S; Cerhan, JR; Severson, RK; Bernstein, L; Hartge, P. 2004 "Epidemiologic evaluation of measurement data in the presence of detection limits." *Environ. Health Perspect.* 112(17):1691-1696.

Maldonado, G. 2008. "Adjusting a relative-risk estimate for study imperfections." *J. Epidemiol. Community Health* 62(7):655-663. doi: 10.1136/jech.2007.063909.

Meliker, JR; Goovaerts, P; Jacquez, GM; Nriagu, JO. 2010. "Incorporating individual-level distributions of exposure error in epidemiologic analyses: An example using arsenic in drinking water and bladder cancer." *Ann. Epidemiol.* 20(10):750-758. doi: 10.1016/j.annepidem.2010.06.012.

Mink, PJ; Goodman, M; Barraj, LM; Imrey, H; Kelsh, MA; Yager, J. 2004. "Evaluation of uncontrolled confounding in studies of environmental exposures and neurobehavioral testing in children." *Epidemiology* 15(4):385-395.

Morgan, MK; Sheldon, LS; Croghan, CW; Jones, PA; Robertson, GL; Chuang, JC; Wilson, NK; Lyu, CW. 2005. "Exposures of preschool children to chlorpyrifos and its degradation product 3,5,6-trichloro-2-pyridinol in their everyday environments." *J. Expo. Anal. Environ. Epidemiol.* 15(4):297-309.

Prueitt, RL; Goodman, JE; Bailey, LA; Rhomberg, LR. 2011. "Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos." *Crit. Rev. Toxicol.* 41(10):822-903.

Rauh, V; Arunajadai, S; Horton, M; Perera, F; Hoepner, L; Barr, DB; Whyatt, R. 2011. "7-Year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environ. Health Perspect.* 119(8):1196-1201.

Rauh, VA; Garfinkel, R; Perera, FP; Andrews, HF; Hoepner, L; Barr, DB; Whitehead, R; Tang, D; Whyatt, RW. 2006. "Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children." *Pediatrics* 118(6):e1845-e1859.

Rauh, VA; Whyatt, RM; Garfinkel, R; Andrews, H; Hoepner, L; Reyes, A; Diaz, D; Camann, D; Perera, FP. 2004. "Developmental effects of exposure to environmental tobacco smoke and material hardship among inner-city children." *Neurotoxicol. Teratol.* 26(3):373-385.

Rothman, KJ; Greenland, S; Lash, TL. 2008. *Modern Epidemiology (Third Edition).* Lippincott Williams & Wilkins, Philadelphia, PA. 758p.

US EPA. 2012. "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Chlorpyrifos Health Effects: Minutes of the FIFRA Science Advisory Panel Meeting held on April 10-12, 2012." FIFRA Scientific Advisory Panel. SAP Minutes No. 2012-04. 108p.

US EPA. 2008. "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: The Agency's Evaluation of the Toxicity Profile of Chlorpyrifos: Minutes of the FIFRA Science Advisory Panel Meeting held on September 16-18, 2008." FIFRA Scientific Advisory Panel. SAP Minutes No. 2008-04. 80p., December 17. Accessed at http://www.epa.gov/scipoly/sap/meetings/2008/091608_mtg.htm.

US EPA. 2010b. "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Draft Framework and Case Studies on Atrazine, Human Incidents, and the Agricultural Health Study: Incorporation of Epidemiology and Human Incident Data into Human Health Risk Assessment." FIFRA Scientific Advisory Panel (SAP). SAP Minutes No. 2010-03. FIFRA Scientific Advisory Panel Meeting, held February 2-4 at the Environmental Protection Agency Conference Center, Arlington, Virginia. 87p., February.

US EPA. 2010a. "Framework for Incorporating Human Epidemiologic & Incident Data in Health Risk Assessment (Draft)." Office of Pesticide Programs. 68p., January 7.

US EPA. 2014. "Chlorpyrifos: Revised Human Health Risk Assessment for Registration Review." Office of Chemical Safety and Pollution Prevention. EPA-HQ-OPP-2008-0850-0195. 531p., December 29.

Whyatt, RM; Camann, D; Perera, FP; Rauh, VA; Tang, D; Kinney, PL; Garfinkel, R; Andrews, H; Hoepner, L; Barr, DB. 2005. "Biomarkers in assessing residential insecticide exposures during pregnancy and effects on fetal growth." *Toxicol. Appl. Pharmacol.* 206(2):246-254.

Whyatt, RM; Rauh, V; Barr, DB; Camann, DE; Andrews, HF; Garfinkel, R; Hoepner, LA; Diaz, D; Dietrich, J; Reyes, A; Tang, D; Kinney, PL; Perera, FP. 2004. "Prenatal insecticide exposures and birth weight and length among an urban minority cohort." *Environ. Health Perspect.* 112(10):1125-1132.

Table 1 Columbia Study Subjects Defined as Developmentally Delayed Based on MDI and PDI Scores Measured at 36 Months of Age in Each of Four Chlorpyrifos Exposure Groups Designated by Rauh *et al.* (2006)[a] These data were used in the sensitivity analyses summarized in Tables 2 and 3.

Outcome		Group 1	Group 2	Group 3	Group 4
		< LOD (n = 80)	1st Tertile > LOD (n = 65)	2nd Tertile > LOD (n = 38 or 39)[b]	3rd Tertile > LOD (n = 45 or 44)[b]
PDI	Psychomotor delay	7	3	3	11
	No delay	73	62	35	34
MDI	Mental delay	30	14	11	20
	No delay	50	51	28	24

Notes:

LOD = Limit of Detection; MDI = Mental Development Index; PDI = Psychomotor Development Index.

(a) Table adapted from US EPA (2014, Appendix 2, Attachment 1). Psychomotor and mental delay defined as PDI and MDI scores ≤ 85 by Rauh *et al.* (2006).

(b) The number of subjects in the two highest tertiles of exposure was inconsistent between the two outcomes.

Table 2 Odds Ratio for "Mental Delay" Calculated for Various Chlorpyrifos Exposure Categorization Schemes (Rauh *et al.*, 2006)[a] Rauh *et al.* (2006) presented results only for the exposure categorization scheme that maximized associations with mental delay. As shown here, associations calculated using a variety of other categorization schemes demonstrate that their choice of defining dichotomous exposure groups likely biased results away from the null.

Exposure Categorization	Group 1	Group 2	Group 3	Group 4
	< LOD	1st Tertile > LOD	2nd Tertile > LOD	3rd Tertile > LOD
Groups 2, 3, 4 (high) *vs.* Group 1 (low)	Reference	0.73 (0.41, 1.29)		
Groups 3, 4 (high) *vs.* Groups 1, 2 (low)	Reference		1.37 (0.78, 2.42)	
Group 4 (high) *vs.* Groups 1, 2, 3 (low)[b]	Reference			**1.95 (1.00, 3.83)**
Groups 2, 3, 4 Individually (high) *vs.* Group 1 (low)	Reference	**0.46 (0.22, 0.96)**	0.65 (0.29, 1.50)	1.39 (0.66, 2.93)
Trend Across Four Dose Groups	Linear trend p = 0.49 OR = 1.09 (0.85, 1.40) for each increase in category			

Notes:
LOD = Limit of Detection; OR = Odds Ratio.
(a) Crude ORs for "Mental Delay" (*i.e.*, MDI score ≤ 85) were calculated using counts of subjects in Table 1. Statistically signification associations at a 95% confidence level are highlighted in bold.
(b) The categorization scheme used by Rauh *et al.* (2006).

Table 3 Odds Ratio for "Psychomotor Delay" Calculated for Various Chlorpyrifos Exposure Categorization Schemes (Rauh *et al.*, 2006)[a] Rauh *et al.* (2006) presented results only for the exposure categorization scheme that maximized associations with psychomotor delay. As shown here, associations calculated using a variety of other categorization schemes demonstrate that the choice of defining dichotomous exposure groups likely biased results away from the null.

Exposure Categorization	Group 1	Group 2	Group 3	Group 4
	< LOD	1st Tertile > LOD	2nd Tertile > LOD	3rd Tertile > LOD
Groups 2, 3, 4 (high) *vs.* Group 1 (low)	Reference	1.35 (0.54, 3.41)		
Groups 3, 4 (high) *vs.* Groups 1, 2 (low)	Reference		**2.73 (1.16, 6.48)**	
Group 4 (high) *vs.* Groups 1, 2, 3 (low)[b]	Reference			**4.23 (1.75, 10.2)**
Groups 2, 3, 4 Individually (high) *vs.* Group 1 (low)	Reference	0.50 (0.13, 2.03)	0.89 (0.22, 3.67)	**3.37 (1.20, 9.46)**
Trend Across Four Dose Groups	Linear trend p = 0.016 **OR = 1.59 (1.01, 2.31) for each increase in category**			

Notes:
LOD = Limit of Detection; OR = Odds Ratio.
(a) Crude ORs for "Psychomotor Delay" (*i.e.* PDI score ≤ 85) were calculated using counts of subjects in Table 1. Statistically signification associations at a 95% confidence level are highlighted in bold.
(b) The categorization scheme used by Rauh *et al.* (2006).

Table 4 Sensitivity Analysis of Outcome Misclassification for CBCL-Derived Outcomes in Rauh *et al.* (2006)[a] Mothers of children with high prenatal exposure to chlorpyrifos may be more likely to report symptoms on the subjective CBCL survey. We recalculated ORs for CBCL-derived outcomes based on a scenario in which 1, 2 or 3 children in the high exposure category were misclassified as having each outcome. Our results show that misclassification of a small number of exposed subjects strongly attenuates the magnitude of associations.

Outcome	Reported ORs		Crude ORs Calculated Assuming 1, 2, or 3 Exposed Subjects Misclassified with a "Problem"		
	Crude	Adjusted	1 Subject	2 Subjects	3 Subjects
Attention Problems	**11.31 (1.75, 120.89)**	**11.26 (1.79, 70.99)**	**8.83 (1.20, 99.31)**	6.46 (0.71, 78.73)	4.20 (0.29, 59.07)
ADHD Problems	**5.59 (1.14, 29.20)**	**6.50 (1.09, 38.69)**	4.36 (0.77, 24.26)	3.20 (0.45, 19.54)	2.08 (0.18, 15.01)
PDD Problems	2.45 (0.50, 10.15)	**5.39 (1.21, 24.11)**	1.80 (0.29, 8.26)	2.05 (0.18, 14.76)	1.00 (0.20, 10.44)

Notes:
ADHD = Attention Deficit Hyperactivity Disorder; CBCL = Child Behavior Checklist; OR = Odds Ratio; PDD = Pervasive Developmental Disorder.
(a) The reported crude and adjusted ORs for behavioral outcomes associated with high vs. low chlorpyrifos exposure were presented in Rauh *et al.* (2006). The recalculated crude ORs were determined based on an assumption that one or more children in the high exposure category were misclassified as having a "CBCL-related problem." Statistically signification associations at a 95% confidence level are highlighted in bold.

Dow AgroSciences

9330 Zionsville Road Indianapolis, IN 46268

www.dowagro.com

Bldg 308/2E
November 12, 2013

Dr. Steve Bradbury
Office of Pesticide Programs (7504P)
U. S. Environmental Protection Agency
One Potomac Yard
2777 S. Crystal Drive
Arlington, VA 22202

CHLORPYRIFOS - A SCIENTIFIC PERSPECTIVE ON THE RELIABILITY AND UTILITY OF INFORMATION FROM EPIDEMIOLOGY STUDIES

Dear Dr. Bradbury,

It is our understanding that the Office of Pesticide Programs (OPP) is preparing a revised human health risk assessment for chlorpyrifos and that this risk assessment could be available for public comment in 2014. As OPP health scientists work toward completion of this task, existing epidemiology data must be considered in addition to many other complex, technical toxicology datasets that have been developed for chlorpyrifos over a number of years.

In considering epidemiology data, it is important that the data be critically evaluated to determine whether they are sufficient to inform risk assessment. Though robust epidemiologic studies have the potential to be useful, preliminary and incomplete studies are often used to fuel sensationalized media messages that may lead to unwarranted public fear and confusion. A recent example of how certain epidemiology studies have been misrepresented and pushed beyond what can reasonably be concluded for chlorpyrifos is the September 3, 2013 article by the Pesticide Action Network in Mother Earth News, entitled, "*Pesticides in Food are Keeping Children From Learning.*"

EPA's Scientific Advisory Panel noted in 2010 when commenting on the Office of Pesticide Programs (OPP) proposed framework to incorporate human studies in its health risk assessments: "*like all information considered in risk assessments, the quality and reliability of the information provided by epidemiologic studies needs to be closely scrutinized.*" Thus, Dow AgroSciences (DAS) would like to take this opportunity to provide you with a scientific perspective on the reliability and utility of the information provided within the epidemiology studies available for chlorpyrifos. We have engaged scientists within DAS as well as qualified external scientific and regulatory consultants to prepare the attached white paper. Within this paper, a number of citations are provided to support the points, which we hope will be useful to the assessment team. We would be happy to answer questions on the document or provide any additional clarifications or information you or your scientific staff may need.

Respectfully,

Bruce Houtman
Leader
U.S. Regulatory & Government Affairs-Crop Protection.

cc: Rick Keigwin, USEPA
Jack Housenger, USEPA
Joel Wolf, USEPA
John Cuffe, Dow AgroSciences
Darin Lickfeldt, Dow AgroSciences

Epidemiology Studies Pertaining to Chlorpyrifos Exposures: Considerations of Reliability and Utility

Debra Edwards, Ph.D., Independent Consultant
Daland Juberg, Ph.D., ATS, Dow AgroSciences
Carol Burns, Ph.D., MPH, The Dow Chemical Company
Julie Goodman, Ph.D., DABT, Gradient
Abby Li, Ph.D., Exponent
Michael Bartels, Ph.D., The Dow Chemical Company
Darin Lickfeldt, Ph.D., Dow AgroSciences

Summary

The utility and application of epidemiology data in risk assessment and regulatory decision-making has received considerable attention in recent years and continues to be vetted by the U.S. Environmental Protection Agency (USEPA) when evaluating chemical risk to human populations (*e.g.*, SAP 2010). For the insecticide chlorpyrifos, there exists a growing number of studies that may inform the risk assessment for this chemical, although one cohort investigated by Columbia University is being considered by the USEPA as providing evidence for the relationship between chlorpyrifos exposure and children's development and cognitive function (*i.e.*, Rauh *et al.*, 2006, Rauh *et al.*, 2011; Whyatt *et al.*, 2004). A critical analysis of published information from varying scientific disciplines and perspectives reveals that findings from this singular cohort have limitations, including reliability of reported results, exposure to other risk factors, lack of reproducibility of findings in other studies, and incompatibility with the voluminous toxicology database for chlorpyrifos (Eaton *et al.*, 2008; Prueitt *et al.*, 2011; Li *et al.*, 2012; Burns *et al.*, 2013). In fact, the European Food Safety Authority (EFSA) recently concluded in its review on epidemiological studies linking exposure to pesticides and health effects that there is "no evidence" to suggest an association between pesticide exposure and neurodevelopmental related outcomes, due to a number of deficiencies in the available data (Ntzani *et al.*, 2013). This EFSA review included neurodevelopment/IQ studies on chlorpyrifos that were published in 2006 and later. The totality of problems relating to the reliability of the reported findings on the Columbia cohort renders the study inappropriate for risk assessment. The study is not useful for informing the question of whether neurodevelopmental effects occur at exposure levels lower than those associated with acetylcholinesterase inhibition or for "bounding" dose-response estimates from animal studies.

The following are key reasons for our position:

1. The analytical method used in the Columbia study has not been validated at the low concentrations reported in maternal/cord blood from Columbia study subjects. Also, the exposure assessment picture is incomplete for chlorpyrifos (and other chemical exposures), given that the reported effects on neurodevelopment (which is a continuous process throughout pregnancy and during early infant and childhood years) are predicated upon a single exposure measurement (snapshot in time).

2. There is a lack of credible scientific evidence to support the biological plausibility of the Columbia cohort findings at the exposures reported. The animal database for chlorpyrifos is one of the most robust across all chemistries and it is not possible to explain adverse neurodevelopmental effects occurring below the threshold point of departure for acetylcholinesterase without significant speculation and conjecture.

3. Chlorpyrifos cannot reliably be deemed as a causal factor for the reported neurodevelopmental effects. There are credible alternative explanations for the observed effects. In particular, influences of other chemical and nonchemical stressors could contribute to or account for the reported associations of impaired neurodevelopment in the Columbia study. Also many of the outcomes reported may be chance occurrences, given the methods used to assess cognitive function.

4. When considering the specificity of biomarkers for chlorpyrifos exposure and lack of concordance across reported outcomes, the results reported in the Columbia study have not, in effect, been replicated or confirmed in other independent studies.

5. Although sufficient information is available to identify the serious limitations with the Columbia study, nonetheless the data on which the Columbia study findings are based have not been shared in a public format for further independent evaluation by either government agencies or other interested investigators. This is not only counter to the principles of transparency in federally funded research, but precludes further scientific analysis which could ultimately assist in rendering a more informed and objective analysis of the data.

1. The analytical method used in the Columbia study has not been validated at the low concentrations reported in blood:

Because a majority of samples from the Columbia cohort are below the validated limit of detection (LOD) for chlorpyrifos analysis in plasma/serum, any conclusions regarding outcomes and associated exposure based on chlorpyrifos levels < 15 picograms (pg)/gram (g) are not reliable. This impacts the reliability of the correlation analyses between exposures and most outcomes since the study dichotomized exposure as "low" and "high" using 6.17 pg/g as its cut point (Whyatt *et al.*, 2004; Rauh et al., 2006). When all values are used in linear analyses, such as when evaluating IQ (Rauh *et al.*, 2011), there may be substantial misclassification of exposure.

It is a basic foundation of the scientific process that researchers must show that a quantitative exposure measurement is accurate, precise, and reproducible across the range of values determined within a study. For example, the USEPA method validation guidelines call for replicate determinations of analyte recovery from a given matrix (substrate) down to the stated LOD (USEPA 1998). However, this has not occurred within the Columbia study. There were no data generated during validation of the plasma/serum analysis method (Barr *et al.*, 2002) or during the subsequent analysis of the Columbia cohort samples to show that chlorpyrifos levels could be accurately measured in plasma/serum matrix down to the stated LOD of 0.5-1 pg/g (note: authors

use LOD term when discussing limit of quantitation). The lowest concentration for which analyte recovery in plasma/serum was determined using this method was 15 pg/g. This is a critical point, as more than 80% of the Columbia subjects had levels below this validation level. Further, there was no evaluation of possible sample contamination during blood collection in the hospital, processing to plasma, or during shipment to the CDC. Analysis of sample integrity is a critical parameter of all biomonitoring studies, especially those at the trace levels reported for this cohort. Note that Barr *et al.* (2002) also reported background chlorpyrifos levels of 9 pg/g in control serum samples, 50% higher than the "high" exposure Columbia cohort criteria, the source of which was never determined. Finally, the CDC has stated that chlorpyrifos blood measurements from the 2003-2004 NHANES survey will not be released because the CDC lab was not able to meet its own QC/QA (Quality Control/Quality Assurance) criteria for the assay (personal communication, CDC). This is the same method used in the Columbia study, so any decisions based on the use of such methodology should be highly scrutinized.

2. There is not a supportable basis for biological plausibility:

Based on a consideration of all available data, acetylcholinesterase inhibition (AChEI) data remain the most sensitive and robust source of dose-response data for deriving points of departure for chlorpyrifos (EPA SAP 2012, p. 28). Currently, there is no established biological mode of action to explain the potential neurodevelopmental effects reported in the Columbia study (EPA SAP 2012, p.30). Based on reliable animal experimentation, neurodevelopmental and/or behavioral effects have been reported at higher exposures, *i.e.*, at or above 1.0 mg/kg body weight per day which appears to be a threshold below which neurodevelopmental effects have not been reported (Li *et al.*, 2012; Maurissen *et al.*, 2000; EPA SAP 2012, p. 36). As noted by the SAP in 2012 (p.36), " . . . *effects of CPF at 1 mg/kg are difficult to interpret because of methodological limitations, inconsistencies, and variation in study design, sometimes lack of control for litter effects, oversampling issues, behavioral methods used, and lack of dose-response findings. At doses exceeding 1 mg/kg, the data show somewhat more consistency, but even here, dose response experiments are the exception."*

The threshold of 1 mg/kg/day is 30-fold higher than the threshold of 0.03 mg/kg/day for the most sensitive red blood cell (RBC) AChEI metric (USEPA 2011 p. 25). The quantitative dose response data for AChEI are especially robust and comprehensive, and include data at the time of peak effect from different ages including rat fetus, young pups, and pregnant and non-pregnant adult rats (Mattsson *et al.,* 2000; Maurissen *et al.,* 2000, Marty and Andrus, 2010).

This threshold of 1 mg/kg/day is also at least 5,000 times higher than estimated exposures from the Columbia study. Lowe *et al.* (2009) estimated prenatal chlorpyrifos exposure levels in the Columbia study to be 0.15 ug/kg/day based on mean maternal and cord blood concentrations reported by Whyatt *et al.,* 2005. These dose levels do not produce RBC AChEI in humans (Garabrant *et al.*, 2009; Farahat *et al.*, 2011). Using a biomonitoring equivalent approach, all blood concentrations for the Columbia study subjects, as well as those of other epidemiology cohorts, were predicted to be well below the level of RBC AChE inhibition (see Attachment A).

One hypothesis that has been put forth is that the neurodevelopmental effects at these low dose levels are a result of hypothetical non-cholinergic modes of action reported in the animal literature. However, as discussed previously, the animal literature indicates that dose levels that cause adverse neurodevelopmental outcomes only occur at exposures that inhibit AChE in pregnant rats or offspring (Li *et al.*, 2012; Prueitt *et al.*, 2011). Thus, the scientific data support that neurodevelopmental effects attributed to chlorpyrifos in the rodent occur at doses at or above 1 mg/kg/day. Based on the animal model, it would take significant conjecture and speculation to conclude that altered neurodevelopment in humans resulting from non-cholinergic pathway perturbations would occur at doses lower than those associated with AChE inhibition.

3. There are credible alternative explanations for the neurodevelopmental effects observed in the Columbia study:

Chlorpyrifos cannot reliably be deemed a causal factor for the neurodevelopmental effects reported by Columbia University. Alternative explanations are credible and present themselves logically when considering exposure measurement error (as discussed under point #1 and below), the incompatibility with the rodent model (as discussed under point #2), methodological issues with analysis of data (as discussed below and in Attachment B), exposure to other toxic chemicals, including neurotoxicants, and the published literature documenting that children who grow up in poverty or low income households have difficulties with neurocognitive function. In commenting on the Columbia University study findings, an independent group of experts acknowledged, "*The authors attempted to control for confounding factors, including other known neurodevelopmental risk factors in this inner-city cohort, such as maternal perinatal smoking and alcohol; nevertheless, it is difficult to dismiss the contribution of these and perhaps other confounding factors*" (Eaton *et al.,* 2008).

The mothers and children within the Columbia study who had measurable exposures to chlorpyrifos were also exposed to other chemicals that have the potential to subtly or profoundly affect child neurodevelopment. For example, the Columbia cohort was exposed to polycyclic aromatic hydrocarbons (PAHs), which were reported to be associated with neurodevelopmental effects in these same children (Perera *et al.,* 2006). Also, lead levels are an important variable in the Columbia study, especially for low-income families living at or near the poverty level. Blood lead levels have consistently been correlated with IQ loss (Healey *et al.,* 2010), as well as achievement and behavioral deficits (Chandramouli *et al.,* 2009). Lead levels were not properly controlled in the Columbia study (Rauh *et al.*, 2006; 2011) for the entire sample, and it is plausible that associations between chlorpyrifos and neurodevelopmental effects could be partially or wholly attributable to lead (see Attachment B).

Poverty and pesticide exposure are highly correlated. Mothers and children who live in crowded, substandard housing are more likely to encounter exposure to multiple and heavily applied pesticides, both legal and illegal (*e.g.,* Morbidity and Mortality Weekly,1997). Indeed, a survey of the Columbia cohort indicated that pesticide use was frequent (Whyatt *et al.,* 2002). It has been proposed that insecticide exposure (regardless of the chemical) may be a marker for insect infestation (and other related factors) and may not itself be the causal agent driving the neurodevelopmental results (Burns *et al.,* 2013). In fact, the Columbia authors reported an

association with the Bayley Scales of Infant Development (BSID) and piperonyl butoxide, a synergist used with pyrethrins and synthetic pyrethroid insecticides, which replaced the use of chlorpyrifos in residential settings after 2001.

It is well documented in the literature that social hardships related to poverty and maternal depression can affect scores on intelligence tests and other measures of cognitive ability (*e.g.,* Luby *et al.,* 2013; Duncan and Brooks-Gunn, 1997; Feinstein, 2003; Canadian Paediatric Society, 2004; center on the Developing Child at Harvard University, 2009). In the Columbia study (Rauh *et al.,* 2006), neurodevelopmental effects were only observed at 3 years of age, not before. As the children age from birth to 3 years, there are a number of other well-known nonchemical risk factors that affect brain development. Efforts were made by the Columbia study investigators to account for certain risk factors, including the influence of race/ethnicity, gestational age, maternal education and maternal IQ (albeit, there were missing IQ data for several dozen women in the study). Observational data on the quality of the home care-taking environment were also considered, but it is unclear to what extent key risk factors were addressed. For example, information was collected on mothers' feelings and state of mind but there is no indication that these potential risk factors were explicitly addressed. Interestingly, maternal depression was a significant factor for influencing childhood behavior, as modeled by the UC Berkeley investigators (Eskenazi *et al.*, 2007; Marks *et al.*, 2010).

Despite efforts made by the Columbia study investigators to account for other risk factors, the influences of other chemical and nonchemical stressors which could contribute to or account for the observed associations of impaired neurodevelopment cannot easily be attributed to the independent effects of a single chemical (*i.e.,* chlorpyrifos) in the multi-chemical exposure scenario experienced by the Columbia cohort, particularly spanning a multi-year period that encompasses an important period of sequential neurodevelopment (*e.g.,* SAP, 2012; Eaton *et al.*, 2008). Any inferences based on the Columbia study regarding the degree to which chlorpyrifos contributes to the measured outcomes cannot be separated easily from other risk factors, and thus the study cannot be used to reliably address the question of whether chlorpyrifos can cause neurodevelopmental effects at the exposure levels reported. Although it is legitimate for academic scientists to propose and investigate hypotheses, the Columbia study cannot serve as a reliable basis for addressing key questions regarding a single chemical in a regulatory risk assessment.

Another significant weakness of the Columbia study relates to the exposure data, *i.e.,* measurements of chlorpyrifos do not reflect exposure over time. Evaluations of neurodevelopmental scores/function on the cohort continued into childhood and early adolescence, which is well beyond the single snapshot in time of chlorpyrifos measurement. This is especially pertinent since neurodevelopment occurs both prenatally and postnatally, essentially a continuous process throughout early infant and childhood years (Selevan *et al.,* 2000). The Columbia University study focused exclusively on prenatal exposure as measured by cord/maternal blood measures of chlorpyrifos within two days of birth. Furthermore, the maternal and cord blood measurements represent a single sample, or snapshot (*i.e.,* only one point in time), collected for convenience (at birth) and with no information regarding the chlorpyrifos home application. Given the rapid metabolism of chlorpyrifos in humans and subsequent short residence time in the body, a single sample obtained at the time of delivery or shortly after would have little relationship or meaning to exposure levels that were present during

most of the pregnancy (or thereafter). Also, as discussed earlier in this paper (Point #1), the chlorpyrifos blood measurements cannot be deemed accurate. The inadequate investigation on exposures to either chlorpyrifos or other pesticides and chemicals (*e.g.,* polycyclic aromatic hydrocarbons, lead, *etc.*) results in an incomplete exposure picture.

Lastly, there are issues regarding how cognitive testing was assessed in the Columbia cohort (Rauh *et al.,* 2006) raising the question of whether the reported associations with chlorpyrifos are real or not. Briefly, the cohort was inappropriately dichotomized into two exposure groups and use of a cut-off of a standard score of "85" for BSID scores to denote children as "High Risk" is an arbitrary decision. Also, the "multiple simultaneous" comparisons in Rauh *et al.* 2006 and 2011 can lead to chance errors (see Attachment B for more details).

4. Adverse results reported in the Colombia study are not found in other populations:

An important aspect of determining the validity of an epidemiology study is whether findings can be reproduced; that is, associations between similar outcomes and exposures to the chemical of interest should be found in different populations. Chlorpyrifos is measured in different ways across studies, with some measuring chlorpyrifos itself, and others measuring other biomarkers that represent exposure to chlorpyrifos and other chemicals, such that exposure to chlorpyrifos itself cannot be teased out. The order of reliability of biomarkers is as follows: chlorpyrifos > 3,5,6-trichloro-2-pyridinol (TCPy) > diethylphosphates (DEPs). The metabolite DEP can reflect exposure to pesticides other than chlorpyrifos. TCPy and DEPs in urine can also result from exposure to these OP metabolites in food or the environment rather than to chlorpyrifos or other OPs.

When considering the order of reliability of biomarkers, the results are not consistent across the existing epidemiology studies. (See Attachment C of this document for details). Specifically, studies at Mount Sinai Hospital and the University of California (UC) at Berkeley do not confirm the results reported by the Columbia University researchers (Eaton *et al.,* 2008; Prueitt *et al.*, 2011; Li *et al.*, 2012; Burns *et al.* 2013). The Mount Sinai (Berkowitz *et al.*, 2004; Engel *et al.*, 2011) and UC Berkeley (Eskenazi *et al.,* 2004: Eskenazi e*t al.*, 2007; Marks *et al.*, 2010; Bouchard *et al.*, 2011) studies report some neurodevelopmental effects associated only with DEP, a less specific biomarker. Outcomes associated with the more specific biomarker, TCPy, are either not tested or not analyzed. Further, there are two new epidemiology studies that have not observed consistent associations with birth weight or developmental outcomes. Two of these studies measured chlorpyrifos or TCPy with reported exposure levels higher than (China cohort; Wickerham *et al.,* 2012) or comparable to (Mexico City cohort; Fortenberry *et al.,* 2013) the Columbia or UC Berkeley studies. Although these studies did not investigate all of the outcomes measured in the Columbia study, findings for endpoints that were evaluated do not confirm findings from the Columbia study (Rauh *et al.,* 2006 and 2011; Whyatt *et al.*, 2004). Based on epidemiological data published through 2007, Eaton *et al.* (2008) also concluded that there were no consistent associations observed when neurodevelopmental outcomes of the Columbia, Mount Sinai, and Berkley studies were compared.

5. Data access has not been provided:

Transparency and documentation of the decision process are at the core of a credible risk assessment. EPA's Office of Pesticide Programs has a long history of transparency as well as data disclosure in risk assessments to ensure the credibility of its registration and reevaluation decisions. OPP's transparency in its risk assessments and decision-making adheres to President Barack Obama's January 21, 2009 Memorandum to Heads of Executive Departments and Agencies on "Transparency and Open Government" (Obama, 2009). Given the concerns about the reliability of exposure assessment in the Columbia study, there would be value in accessing the data for the purposes of exposure (dose) reconstruction and review of the health effect analyses. Similarly, the UC Berkeley and Mount Sinai studies (Eskenazi *et al.*, 2007; Marks *et al.*, 2010; Bouchard *et al.*, 2011; Engel *et al.*, 2011) conducted no health analyses using chlorpyrifos in cord blood or with urinary TCPy after age two. Access to and independent analyses of these data would also be informative to determine the reliability of the Columbia results.

OPP has indicated publicly, "*The studies that are the most relevant and informative to risk assessment are those that clearly and fully describe study design, conduct and methods, as well as providing access to the underlying data*" (http://www.epa.gov/pesticides/science/literature-studies.html). OPP has considered the Columbia study, which is a federally funded study, as a source of data intended for consideration in its chlorpyrifos risk assessment (see SAP 2008; 2012). However, though DowAgroSciences has made repeated requests through the Freedom of Information Act (FOIA) to the Agency to obtain the Columbia data, there is a restriction placed on data access by the authors. Thus, independent verification of the analyses (including an evaluation of different cutoff points for exposure and BSID outcomes) and the ability to answer specific questions regarding the Columbia study (*e.g.*, other risk factors) are not possible by DowAgroSciences, EPA scientists, or the public. This lack of access is counter to the recent February 22, 2013 John P. Holdren Memorandum to Heads of Executive Departments and Agencies on "Increasing Access to the Results of Federally Funded Scientific Research" (Holdren, 2013). Thus, any significant cited line of evidence in OPP's chlorpyrifos risk assessment should be based on accessible data sets that allow for independent analysis and verification of conclusions. We acknowledge the importance of protecting the privacy of the subjects in epidemiology studies, but there are well-recognized and accepted ways to provide data on cohort subjects while protecting the privacy of the subjects. Given the problems and complex issues involved with the chlorpyrifos cohort data, including the type of cognitive assessment used, a more thorough and multidisciplinary scientific review is needed, which provides some access to the data and includes pediatricians, epidemiologists, clinicians and neuropsychologists experienced in evaluation of pediatric cognitive function.

Conclusions:

Despite not having access to the data published by Columbia University, there is sufficient information available to conclude that there are serious limitations that impact their utility and reliability in risk assessment. The Columbia study is not suitable as a basis to "bound" dose-response estimates from animal studies or to inform whether neurodevelopmental effects occur at exposure levels lower than those associated with AChEI. This is because of the difficulty of

disentangling the potential of other chemical and nonchemical stressors to account for or contribute to the observed associations. Further, the analytical method used in collecting plasma biomonitoring data to address the issue of whether any health outcome could potentially occur below exposure levels resulting in AChEI has not been adequately validated. The incompatibility with the rodent model and the lack of biological plausibility for chlorpyrifos causing neurodevelopmental outcomes in children at estimated exposure levels, as well as the lack of consistency with other populations, indicate that any reported statistical associations within the Columbia study are likely due to factors other than chlorpyrifos exposure. The data are not accessible for public viewing and independent analysis, which is counter to the basic tenets of transparency in government-funded research. Confidence in the reliability of the conclusions for risk assessment-based decisions is inseparably tied to transparency and transparency cannot be achieved when data access is denied. Collectively, these represent very compelling and independent bases for precluding the use of the Columbia study data in an evaluation of exposure and effect for chlorpyrifos or for calling into question the robust and comprehensive animal database on pre- and postnatal toxicity and its adequacy to characterize the dose-response curve at lower dose levels for the young.

References Cited:

Barr, D. B., Barr, J. R., Maggio, V. L., Whitehead, R. D., Jr., Sadowski, M. A., *et al.* (2002). A Multi-Analyte Method for the Quantification of Contemporary Pesticides in Human Serum and Plasma Using High-Resolution Mass Spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 778:99-111.

Berkowitz, G.S., Wetmur, J., Birman-Deych, E., Obel, J., Lapinski, R.H., Godbold, J.H., Holzman, I.R., Wolff, M.S. (2004). In utero pesticide exposure, maternal paraoxonase activity, and head circumference. Environ Health Perspect, 112:388–391.

Bouchard M. F., Chevrier J., Harley K.G., Kogut K., Vedar M., Calderon N., Trujillo C., Johnson C., Bradman A., Barr D.B., Eskenazi B. (2011). Prenatal Exposure to Organophosphate Pesticides and IQ in 7-Year Old Children. Environ Health Perspect. 119:1189-1195.

Burns, C. J., McIntosh, L. J., Mink, P. J., Jurek, A. M., Li, A. A. (2013)., Pesticide Exposure and Neurodevelopmental Outcomes: Review of the Epidemiologic and Animal Studies. J Toxicol Environ Health B Crit Rev 16:127-283.

Canadian Paediatric Society. (2004). Maternal depression and child development. Paediatr. Child Health 9(8):575-583.

Center on the Developing Child at Harvard University (2009). Maternal Depression Can Undermine the Development of Young Children. Working Paper No. 8, http:/www.developingchild.harvard.edu.

Chandramouli, K., Steer, C. D., Ellis, M., and Emond, A. M. (2009). Effects of early childhood lead exposure on academic performance and behaviour of school age children. *Archives of Disease in Childhood*, 2009. 94:844-8.

Duncan, G.J. and J. Brooks-Gunn (eds) (1997), Consequences of growing up poor. New York, N.Y.: Russell Sage Foundation.

Eaton D.L., Daroff R.B., Autrup H., Bridges J., Buffler P., Costa L.G., Coyle J., McKhann G., Mobley W.C., Nadel L., Neubert D., Schulte-Hermann R., Spencer P.S. (2008). Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. Crit Rev Toxicol. 38 Suppl 2:1-125.

Engel, S.M., Wetmur, J., Chen, J., Zhu, C., Barr, D.B., Canfield, R.L., Wolff, M.S. (2011). Prenatal exposure to organophosphates, paraoxonase 1, and cognitive development in childhood. Environ Health Perspect. 118:1182-8.

Eskenazi, B., Harley, K., Bradman, A., Weltzien, E., Jewell, N. P.,*et al.* (2004). Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. Environ Health Perspect. 112(10):1116-24.

Eskenazi,B., Marks, A.R., Bradman, A., Harley, K., Barr, D.B., Johnson, C., Morga, N., and Jewell N.P. (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. Environ Health Perspect. 115(5):792-798.

Farahat, F.M., Ellison, C.A., Bonner, M.R., McGarrigle, B.P., Crane, A.L., *et al.*, (2011). Biomarkers of chlorpyrifos exposure and effect in Egyptian cotton field workers. Environ. Health Perspect. 119(6):801-6.

Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. Economica,70(277),73-98. doi: 10.1111/1468-0335.t01-1-00272.

Fortenberry, G. Z., Meeker, J. D., Sanchez, B. N., Barr, D. B., Panuwet, P., *et al.* (2013). Urinary 3,5,6-Trichloro-2-Pyridinol (Tcpy) in Pregnant Women from Mexico City: Distribution, Temporal Variability, and Relationship with Child Attention and Hyperactivity. Int J Hyg Environ Health. pii: S1438-4639(13)00112-0. doi: 10.1016/j.ijheh.2013.07.018.

Garabrant, D.H., Aylward, L.L., Berent, S., Chen, Q., Timchalk, C; *et al.* (2009). Cholinesterase inhibition in chlorpyrifos workers: Characterization of biomarkers of exposure and response in relation to urinary TCPy. J. Expo. Sci. Environ. Epidemiol. 19(7):634-42.

Healey, N., Jones−Otazo, H., Walker, M., and Knafla, A. (2010). Toxicological Review and Recommended Toxicological Reference Values for Environmental Lead Exposure in Canada. FINAL REPORT. Prepared under contract to Health Canada. Prepared for the Contaminated Sites Division, Safe Environments Directorate, Healthy Environment and Consumer Safety Branch, Health Canada, Ottawa.

Holdren, John P. (2013), Memorandum for Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research. February 22.

Horton, M. K., Rundle, A., Camann, D. E., Barr, D., Rauh, V. A., *et al.* (2011). Impact of Prenatal Exposure to Piperonyl Butoxide and Permethrin on 36-Month Neurodevelopment. Pediatrics 127:699-706.

Li, A., Lowe, K.A., McIntosh, L.J., Mink, P.J. (2012). Evaluation of epidemiology and animal data for risk assessment: chlorpyrifos developmental and neurobehavioral outcomes. J. Toxicol. Env. Health. Part B. 15:109-184.

Lowe, E. R., Poet, T. S., Rick, D. L., Marty, M. S., Mattsson, J. L., *et al.* (2009). The Effect of Plasma Lipids on the Pharmacokinetics of Chlorpyrifos and the Impact on Interpretation of Blood Biomonitoring Data. Toxicol Sci 108:258-272.

Luby, J. V., Belden, A., Botteron, K., Marrus, N., Harms, M.P. *et al.* (2013). The Effects of Poverty on Childhood Brain Development: The Mediating Effect of Caregiving and Stressful Life Events. O JAMA Pediatr. Published online October 28, 2013. doi:10.1001/jamapediatrics.2013.3139.

Marty, M.S., Andrus, A.K., Bell, M.P., Passage, J.K., et. al. (2012). Cholinesterase inhibition and toxicokinetics in immature and adult rats after acute or repeated exposures to chlorpyrifos or chlorpyrifos-oxon. Reg. Toxicol. Pharmacol. 63:209-224.

Mattsson, J.L., Maurissen, J.P.J., Nolan, R.J., and Brzak, K.A. (2000). Lack of differential sensitivity to cholinesterase inhibition in fetuses and neonates compared to dams treated perinatally with chlorpyrifos. Tox. Sci. 53:438-446.

Maurissen, J.P.J., Hoberman, A.M., Garman, R.H., and Hanley Jr., T.R. (2000). Lack of selective developmental neurotoxicity in rat pups from dams treated by gavage with chlorpyrifos. Tox. Sci. 57:250-263.

Marks, A.R., Harley, K., Bardman, A., Kogut, K., Barr, D.B., Johnson, C., Calderon, N., Eskenazi, B. (2010). Organophosphate pesticide exposure and attention in young Mexican-American children: the CHAMACOS study. Environ. Health Perspect. 118:1768-74.

Morbidity and Mortality Weekly (1997). Poisonings associated with illegal use of aldicarb as a rodenticide - New York City. 1994-1997. 46:961-963.

Ntzani, E.E., Chondrogiorgi, M., Ntritsos, G., Evangelou, E., Tzoulaki, I. (2013). Literature review on epidemiological studies linking exposure to pesticides and health effects. EFSA supporting publication 2013:EN-497, 159 pp.

Perera, F. P., Rauh V., Whyatt, R. M. *et al.* (2006). Effect of prenatal exposure to airborne polycyclic aromatic hydrocarbons on neurodevelopment in the first 3 years of life among inner-city children. Environ. Health Perspect. 114(8):1287-1292.

President Barack Obama (2009). Memorandum for the Heads of Executive Departments and Agencies: Transparency and Open Government. January 21.

Prueitt, R.L., Goodman, J.E., Bailey, L.A., Rhomberg, L.R. (2011). Hypothesis-based weight-of-evidence evaluation of the neurodevelopmental effects of chlorpyrifos. Crit. Rev. Toxicol. 41(10):822-903.

Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L. *et al.* (2006). Impact of Prenatal Chlorpyrifos Exposure on Neurodevelopment in the First 3 Years of Life among Inner-City Children. Pediatrics 118:e1845-1859.

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D.B., and Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ. Health. Perspect. 119(8) 1196-1201.

Selevan, S. G., Kimmel, C. A., Mendola P. (2000). Identifying critical windows of exposure for children's health. Environ. Health Perspect. (June); 108 (Suppl 3): 451-455.

US EPA guidelines for Exposure and Risk Assessment Calculations (1998), Series 875-Occupational and Residential Exposure Test Guidelines: Group B- Post application Exposure Monitoring Test Guidelines, Version 5.4, Working Draft of February 10, 1998, Method Validation criteria pp. C6-7.

US EPA Office of Pesticide Program. (2011). Chlorpyrifos:Preliminary Human Health Risk Assessment for Registration Review. DP No. D388070.

US EPA, FIFRA Scientific Advisory Panel. (2008). A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: The Agency's Evaluation of the Toxicity Profile of Chlorpyrifos: Minutes of the FIFRA Science Advisory Panel Meeting held on September 16-18, 2008. SAP Minutes No. 2008-04. 80p.

US EPA, FIFRA Scientific Advisory Panel. (2010). Incorporation of Epidemiology and Human Incident Data into Human Risk Assessment: Minutes of the FIFRA Science Advisory Panel Meeting held on February 2 - 4, 2010, at http://www.epa.gov/scipoly/sap/meetings/2010/020210meeting.html

US EPA, FIFRA Scientific Advisory Panel. (2012). A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Chlorpyrifos Health Effects: Minutes of the FIFRA Science Advisory Panel Meeting held on April 10-12, 2012. SAP Minutes No. 2012-04. 108p.

Whyatt, R. M., Camann, D. E., Kinney, P. L., *et al.* (2002)., Resdiential pesticide use during pregnancy among a cohort of urban miority women. Environ. Health Perspect. 110:507-514.

Whyatt, R. M., Rauh, V., Barr, D. B., Camann, D. E., Andrews, H. F. *et al.* (2004). Prenatal Insecticide Exposures and Birth Weight and Length among an Urban Minority Cohort. Environ Health Perspect 112:1125-1132.

Whyatt, R. M., Camann, D., Perera, F. P., Rauh, V. A.,Tang, D. *et al.* (2005). Biomarkers in Assessing Residential Insecticide Exposures During Pregnancy and Effects on Fetal Growth. Toxicol Appl Pharmacol 206:246-254.

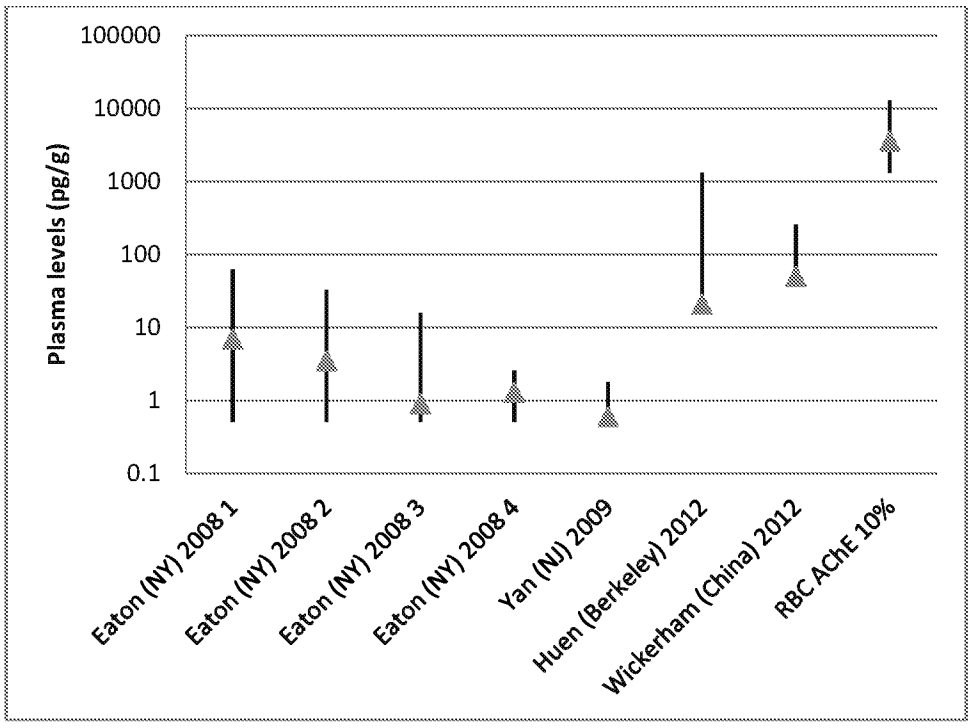
Wickerham, E. L., Lozoff, B., Shao, J., Kaciroti, N., Xia, Y. *et al.* (2012). Reduced Birth Weight in Relation to Pesticide Mixtures Detected in Cord Blood of Full-Term Infants. Environ Int 47:80-85.

Attachment A: Consideration of AChE depression in Columbia study subjects

The purpose of this attachment is to provide additional explanation supporting the prediction, based on the available biomonitoring data, that all blood concentrations for the Columbia study subjects, as well as those of other epidemiology cohorts, were well below the level of RBC AChE inhibition.

It is cumbersome to compare chronic dose levels administered to animals with concentration levels observed in spot samples collected in humans. Biomonitoring equivalents (BE) address this problem. For chlorpyrifos, BE values were developed for blood CPF that are associated with a predicted maximum of 10% inhibition of red blood cell acetylcholinesterase (RBC AChE) using a physiologically based pharmacokinetic/ pharmacodynamic (PBPK/PD) model (Arnold *et al.*, 2013). Inhibition of RBC AChE, while not part of the cholinergic toxicity pathway in the central nervous system, is a conservative marker of inhibition of brain AChE, since inhibition of these blood enzymes occurs pre-systemically in the liver during CPF metabolism, and is the USEPA regulated endpoint.

There are four human studies in which chlorpyrifos was measured in cord blood and or maternal serum (Whyatt *et al.*, 2004; Yan *et al.*, 2009; Huen *et al.*, 2012; Wickerham *et al.*, 2012). The cord blood levels are shown for each study in the graph below. Since the levels declined appreciably over time in the Columbia study, the levels are shown by year of birth (as reported in Whyatt *et al.*, 2004 and Eaton *et al.*, 2009). Also shown is the range of plasma chlorpyrifos concentrations predicted to cause 10% inhibition of RBC AChE (RBC AChE 10%). A few subjects in the UC Berkeley study (Huen *et al.*, 2012) may have had concentrations near the lowest estimate for 10% RBC AChE inhibition. All other study subjects were well below all estimates for RBC AChE inhibition.



AChE: Acetylcholinesterase; Eaton 2008 1 – 4 (birth years 1999, 2000, 2001, 2002)

References:

Arnold, S., Morriss, A., Velovitch, J., Juberg, D., Price, P., *et al.* (2013). Derivation of Biomonitoring Equivalents for Chlorpyrifos Using a Pharmacokinetic/Pharmacodynamic Model of Oral Exposures. *The Toxicologist.* 52nd Society of Toxicology Annual Meeting and ToxExpo. March 10-14, 2013, San Antonio, TX p. 268.

Eaton, D.L., Daroff, R. B., Autrup, H., Bridges, J., Buffler, P. *et al.* (2008). Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. *Crit Rev Toxicol* 38 Suppl 2:1-125.

Huen, K.,Bradman, A., Harley, K., Yousefi, P., Boyd Barr, D., *et al.* (2012). Organophosphate Pesticide Levels in Blood and Urine of Women and Newborns Living in an Agricultural Community. *Environ Res* 117:8-16.

Whyatt R.M., Rauh V., Barr D.B., Camann D.E., Andrews H.F., Garfinkel R., Hoepner L.A., Diaz D., Dietrich J., Reyes A., Tang D., Kinney P.L., Perera F.P. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environ Health Perspect* 112:1125-32.

Wickerham, E. L.;Lozoff, B.;Shao, J.;Kaciroti, N.;Xia, Y., *et al.* (2012). Reduced Birth Weight in Relation to Pesticide Mixtures Detected in Cord Blood of Full-Term Infants. *Environ Int* 47:80-85

Yan, X., Lashley, S., Smulian, J. C., Ananth, C. V., Barr, D. B., Ledoux, T. A., Hore, P., Robson, M. G. (2009). Pesticide Concentrations in Matrices Collected in the Perinatal Period in a Population of Pregnant Women and Newborns in New Jersey, USA, *Human Ecolog Risk Asses* 15:948 - 967

Attachment B: Scientific perspective on specific Columbia study parameters and analyses

The purpose of this attachment is to provide additional points contributed by Dr. Alan S. Kaufman (Clinical Professor of Psychology at the Yale Child Study Center at the Yale University School of Medicine) related to the Columbia study findings (Rauh et al., 2006 and Rauh et al., 2011) and reported interpretation within the broader context of developmental/cognitive testing and assessment. Personal Communication, November 6, 2013.

Columbia Cohort and Lead Exposure

Please comment on exposure of the Columbia cohort to lead and its impact on the exposure-response relationship to chlorpyrifos in the Columbia study?

Lead levels are an important variable in the Columbia study, especially for low-income families living at or near the poverty level. Blood lead levels have consistently been correlated with IQ loss (Healey *et al.*, 2010), as well as achievement and behavioral deficits (Chandramouli *et al.*, 2009). In the Columbia cohort, it was reported "*Lead levels were available, however, for only a subset of children (n = 89). Within this subset, there was no significant relationship between prenatal lead levels and chlorpyrifos levels (r = -0.08; P = .49).*" (Rauh *et al.*, 2006). However, an important question regarding lead that is not addressed in the Columbia study is how did high prenatal lead levels correlate with changes in the Bayley mental and motor development scores? Lead levels were not properly controlled in the Columbia study for the entire sample, and it is plausible that significant relationships could have emerged between lead levels and high chlorpyrifos levels and also between lead levels and Bayley scores. Thus, it is feasible that lead - uncontrolled in this study - played an unknown role in the subjects with "high" chlorpyrifos levels.

Although Rauh *et al.* in the 2006 study neglected to examine the relationship between cord lead and Bayley scores, in their 2011 study, Rauh *et al.* did determine whether cord lead was related to both chlorpyrifos and the WISC-IV scores. Nonetheless, there was cord blood information on too few mothers to be able to control this important variable and blood lead levels of the children in the study were lacking, which should have been done when the children were 1-3 years old.

Columbia Cohort and Cognitive Testing Assessment

Please comment on the use of dichotomized scores on the Bayley Scales of Infant Development (BSID) outcomes.

Regarding the dichotomizing of the Columbia cohort into two portions: Rauh *et al.* (2006) say the issue concerns the dichotomizing of the subjects into two portions. Rauh et al. (2006) indicate: "*The most highly exposed group and the undetectable group had lower mean MDI and PDI scores than did the 2 middle levels. On the basis of these preliminary analyses, and consistent with our previous reports, a dichotomized exposure variable was used, classifying subjects into high exposure (>6.17 pg/g) or lower exposure (≤6.17 pg/g).*" The four groups should have been analyzed separately. There was a "U-shaped" relationship between the Bayley scores and level of chlorpyrifos exposure. The "undetectable" group and the "high exposure"

group both scored lowest on the mental and motor scales of the Bayley. There is no scientific justification for combining the two middle groups with the "undetectable" group. Further, it is inappropriate to examine the Bayley scores for the four groups *before* making the decision of how to combine the data. The Bayley scores are the dependent variable for the study (*i.e.*, the outcome variables). It is not good scientific practice to examine data on the outcome variables before deciding how to analyze the data.

Columbia Cohort and Mental/Psychomotor Performance Across Exposure Groups

Please comment on the mental and psychomotor delay at age 3 when comparing high to low exposure groups within the Columbia cohort.

Regarding the loss of mental function at age 3 years, Rauh *et al.*, (2006) report that "*Highly exposed children (chlorpyrifos levels of ≥ 6.17 pg/g plasma) scored, on average, 6.5 points lower on the Bayley Psychomotor Development Index and 3.3 points lower on the Bayley Mental Development Index at 3 years of age compared with those with lower levels of exposure. Children exposed to higher, compared with lower, chlorpyrifos levels were also significantly more likely to experience Psychomotor Development Index and Mental Development Index delays, attention problems, attention-deficit/hyperactivity disorder problems, and pervasive developmental disorder problems at 3 years of age.*" Thus, the motor and mental results were treated as if they are the same, which they are not. A 3.3 point discrepancy on the mental index is not a meaningful difference, and that difference did not even approach statistical significance at the 0.05 level ($p = 0.155$ in Table 2; Rauh *et al.*, 2006). It is inappropriate to speak of significant mental "delays" as in the Rauh *et al.*, 2006 paper. First, the significance level = 0.048 (Table 2; Rauh *et al.*, 2006) is barely under the $p < 0.05$ guideline. Nonetheless, that probability has no meaning because of the "multiple simultaneous" comparisons in Table 2 (12 comparisons, to be exact - four contrasts at each of three ages). Whenever more than one comparison is made at a time, it is incumbent on the researcher to exercise some type of control over the chance error that inevitably occurs when many comparisons are made at once. Rauh *et al.* made no such control (*e.g.*, demanding that each separate probability must be $p < 0.02$ or $p < 0.01$ to achieve a "family-wise" error rate of 0.05). In short, the value of 0.048 is "not" significant, but likely a result of making so many comparisons such that a few will be "significant" just by chance occurrence. Secondly, the use of a cut-off of a standard score of "85" to denote children as "High Risk" is an arbitrary decision. Scoring below 85 is not a diagnostic category. It is not even a common "cut" score for determining who is at high risk; values below 85 are much more common. Harrison (1990, pp. 53-56), for example, uses cut-off scores of 70, 75, and 80 (but not 85) to illustrate the use of the *Early Screening Profiles* for identifying high risk children between the ages of 2 to 6 years. Changing categories, in any event, is not meaningful. IQ and motor development tests have a built-in standard error of measurement of 3 or 4 points--and that error is even higher when testing very young children tested on tests developed for infants and toddlers. For example, Black and Matula (2000) point out that for the second edition of the Bayley Scales (normed for ages 1-42 months): *"The average standard error of measurement is 5.21 for the Mental Scale and 6.01 for the Motor Scale"* (pp. 68-69); those values are far lower than the values of about 3.00 for Wechsler's IQ scales at ages 3-7 years (Pearson, 2012, Table 4.3; Psychological Corporation, 2003, Table 4.3). Categories such as "High Risk" will change dramatically from Test 1 to Test 2 when the same child is tested twice simply due to errors of

measurement. Whether a child scores 82 or 83 or 84 or 85 or 86 or 87 is just pure chance due to measurement error. A single arbitrary cut-off is inadequate to identify children as normal or with "delays." Such an arbitrary approach takes unfair advantage of the known errors of measurement that characterize even the best measures of mental and motor ability; 85 is just an arbitrary number that may or may not mean delay or high risk.

In the 2011 study, the Rauh *et al.* analysis of chlorpyrifos and WISC-IV scores appears sound. However, the multiple comparisons (as mentioned above) are still an issue. In Table 2, there are five adjusted value comparisons; the authors made no attempt to control for errors that occur when several comparisons are made at once. Consequently, the $p = 0.048$ value for Full Scale IQ is suspect and most likely a chance finding. Nonetheless, the significant finding for working memory is robust and not likely due to chance. However, whenever only one of four mental indexes is found to be significant, such a finding should be replicated with independent samples to verify that it is a "real" relationship between chlorpyrifos and intelligence, not a chance finding.

References:

Black, M. M., & Matula, K. (2000). Essentials of Bayley Scales of Infant Development-II assessment. Hoboken, NJ: Wiley

Chandramouli, K., Steer, C. D., Ellis, M., and Emond, A. M. (2009). Effects of early childhood lead exposure on academic performance and behaviour of school age children. *Archives of Disease in Childhood*, 2009. 94:844-8.

Harrison, P. L. (1990). *Early Screening Profiles manual.* Minneapolis, MN: Pearson.

Healey, N., Jones-Otazo, H., Walker, M., and Knafla, A. (2010). Toxicological Review and Recommended Toxicological Reference Values for Environmental Lead Exposure in Canada. FINAL REPORT. Prepared under contract to Health Canada. Prepared for the Contaminated Sites Division, Safe Environments Directorate, Healthy Environment and Consumer Safety Branch, Health Canada, Ottawa.

Pearson. (2012). *Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition.* San Antonio, TX.

Psychological Corporation. (2003). *WISC-IV technical and interpretive manual.* San Antonio, TX.

Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., et al. (2006). Impact of Prenatal Chlorpyrifos Exposure on Neurodevelopment in the First 3 Years of Life among Inner-City Children. Pediatrics 118:e1845-1859.

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D.B., and Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Env. Hlth. Perspect. 119(8) 1196-1201.

Attachment C: Epidemiology Data for Chlorpyrifos – Considerations of Reproducibility

The purpose of this attachment is to explain the importance of type and specificity of exposure when evaluating evidence for and against causality, namely using chlorpyrifos and/or urinary TCPy compared to less specific metabolites. Demonstration of reproducibility between studies (not within) is critical for risk assessment decisions.

To support causality, associations between exposures to the chemical of interest and health outcomes should be found in different populations. The summary Table 1 presented below, which uses the format of EPA's Table 10 (USEPA 2012, page 59), includes recent data from two new cohorts (China cohort; Wickerham et al., 2012; Mexico City cohort; Fortenberry et al., 2013) that have exposure levels higher than or comparable to the Columbia and University of California (UC) Berkeley studies. Another study conducted in New Jersey (Yan et al., 2009; Barr et al., 2010) did not observe any significant associations for fetal growth; however, in this study, chlorpyrifos was near the limit of detection in most maternal and cord blood samples.
A major challenge in determining reproducibility of the Columbia study is a lack of consistency of exposure metrics. Some other cohort studies measured urinary biomarkers of organophosphate (OP) exposure that have different levels of specificity as a biomarker for chlorpyrifos exposure. The most specific urinary biomarker of chlorpyrifos exposure is 3,5,6-trichloro-2-pyridinol (TCPy), but this also can be a metabolite of chlorpyrifos-methyl. Dialkylphosphates (DAPs) are non-specific biomarkers of OPs that are comprised of diethylphosphates (DEPs) and dimethyl phosphates (DMPs). The DEPs include two metabolites of chlorpyrifos, diazinon and other diethyl-OPs (Table 2). In contrast, the DMPs are not biomarkers for chlorpyrifos but can be biomarkers of many dimethyl-OPs (Table 2). These urinary metabolites of OPs can also be a result of exposure to these metabolites in food or the environment rather than to chlorpyrifos or other OPs. Thus, in evaluating consistency across studies it is necessary to consider the studies in terms of the level of information they provide about chlorpyrifos specifically, as opposed to information about OPs generally, as well as what is meant by "consistent findings" in the context of these different biomarkers of exposure (Li *et al.*, 2012; Mink *et al.*, 2012). The summary Table 1 in this attachment differs from EPA's Table 10 in that associations with DEPs are tabulated instead of those for DAPs.

Points to Note in Reviewing the Table 1:

1. EPA's 2012 SAP clearly stated that total DAPs "are not selective enough to be a useful biomarker for chlorpyrifos" (EPA SAP, P. 58). Chlorpyrifos is the biomarker deemed to be the highest priority because of its specificity (EPA SAP, P. 58). EPA SAP also concluded that "the next biomarker of choice is TCPy, then DETP/DEP in urine," although both are present in the environment as degradates of the active ingredient

(USEPA SAP, p. 58). Thus, this table omits DAPs but includes DEPs. The columns are shaded to reflect the order of priority based on specificity of the biomarker as described above.

2. We define null findings as those with a p value >0.1, non-significant findings as those with a p value>0.5, and present the direction (either as positive for score increased or inverse for score decreased) for findings with a p value<0.1. Values were calculated from the confidence interval when p values were not provided (Altman and Bland, 2011). The table is only a brief synopsis; for more in depth analysis including tables of the magnitude and direction of effects and all statistical testing conducted, we direct the reader to one of several published reviews (Prueitt *et al*, 2011; Li *et al.*, 2012; Mink *et al.*, 2012, Burns *et al.*, 2013).

3. Table 1 only includes associations reported for prenatal exposures (*i.e.* maternal DEPs and TCPy) to be comparable to the Columbia cohort study, which only measured cord/maternal blood at birth.

4. Few findings of the Columbia cohort have been tested by independent investigators using TCPy and/or chlorpyrifos in blood, although these data are available (Huen *et al.* 2012). For example, the UC Berkeley study has not evaluated any neurodevelopmental outcomes using available cord blood chlorpyrifos levels or reported any results using TCPy in children over 2 years of age. It is unknown why the UC Berkeley investigators have not published any health results using these data that would contribute more informed data for use in decision-making relevant to chlorpyrifos.

5. Pervasive Development Disorder (PDD) and Attention-Deficit Hyperactivity Disorder-like behaviors (ADHA) were not clinically diagnosed in the Table 1 studies. Rather, they were based on checklists completed by the mother. Furthermore, maternal depression, a potential confounder of maternal reporting of behavioral problems in children, was not controlled by the Columbia or Mexico City investigators, although it was a significant factor in the UC Berkeley study.

Analysis of summary table

Table 1 is a high level summary of many analyses and purely looks at statistical associations. Even without discussing methodological differences, most of the findings of the Columbia study are not replicated. After age 2 years, it might appear that the UC Berkeley cohort shows limited consistency with the Columbia cohort because of the association with mothers' report of Attention-Deficit Hyperactivity Disorder-like behaviors (ADHD) and DEPs. Unfortunately, the UC Berkeley study did not report any testing for TCPy or chlorpyrifos in older children. Notably, there was no significant overall association found with ADHD and other attention

problems and TCPy in the Mexico cohort. Both the UC Berkeley and Mexico investigators conducted multiple tests for attention and behavioral problems without control for multiple testing. For example, in Table 5 of the Mexico cohort, 27 tests for trend were presented for which two were considered borderline statistically significant ($p<0.1$) while none was statistically significant at the standard p-level (i.e., ($p<0.05$). Overall, since positive results were only reported for DEP and only at age 5 in the UC Berkeley study, and not for TCPy, there is little support for a consistent exposure and ADHD association and less support for any effects attributable to chlorpyrifos.

Both the UC Berkeley and Mount Sinai cohorts show some consistency with Full Scale IQ with DEPs ($0.05<p<0.1$), which are not specific to chlorpyrifos. Again, no testing was reported for TCPy or chlorpyrifos in the other cohorts. Thus, although one could selectively focus on the statistically significant DEPs association from the UC Berkeley study and the findings in the Columbia study as evidence of consistency across cohorts, the null findings with the more specific biomarker TCPy significantly weakens the weight of evidence. The observations in the Columbia study have not been sufficiently tested with chlorpyrifos exposure in other studies to confirm if these are true or false observations. Without robust replication, the Columbia data should not be used for risk assessment.

Table 1. Summary of findings from key epidemiology studies for chlorpyrifos.

	Columbia	Mount Sinai		UC Berkeley			Mexico City	China
Markers of exposure	CPF	TCPy	DEPs	CPF	TCPy	DEPs	TCPy	CPF
Birth Length	Inverse (Null post 2000)	Null	Null	Collected, analysis not available	Null	Null	N.I.	N.I.
Birth Weight	Inverse (Null post 2000)	Null	Inverse	Collected, analysis not available	Null	Null	N.I.	Null
Bayley Scores 12 months (MDI/PDI)	Null/Null	Collected, analysis not available	Null/null	Collected, analysis not available	Null/Null	Null/Null	N.I.	N.I.
Bayley Scores 24 months (MDI/PDI)	Null/Null	Collected, analysis not available	Null/Null	Collected, analysis not available	Null/Null	Null/Null	N.I.	N.I.
Bayley Scores 36 months (MDI/PDI)	Inverse/ Inverse	Not tested	Not tested	Not tested	Not tested	Not tested	N.I.	N.I.
Pervasive Development Disorder (PDD)	Positive (36 mo)	Not tested	Not tested	Collected, analysis not available	Null (24 mo)	Null (24 mo)	N.I.	N.I.
ADHD/ attention and behavior problems ages 2 - 7 years	Positive (36 mo)	Not tested	Not tested	Collected, analysis not available	Null (24 mo)	Null (3.5 yr) Positive (5 yr)[1]	Null[2] (6-11 yr)	N.I.
Mental Development (WISC-IV, age 7 – 9 years)	Inverse (Full-scale IQ and Working Memory); Null (Others)	Collected, analysis not available	Inverse (NS) FSIQ, Working memory	Collected, analysis not available	Collected, analysis not available	Inverse (FSIQ) Null (working memory)	N.I.	N.I.

1. P< 0.1 for a single test at 3.5 years. All testing for other attention and behavioral problems at age 5 years were not statistically significant (Marks et al., 2010).
2. P < 0.1 for a single ADHD index in boys and Hit RT block change in all subjects. All testing for other models for changes in psychometric assessment scores were not statistically significant (Fortenberry, et al., 2013).
NS=Not statistically significant, [0.1>p>0.05]. p values calculated from confidence interval using following equation: SE = (u - l)/(2 x 1.96); z = Est/SE; P = exp(-0.717 x z - 0.416 x z^2) (Altman and Bland, 2011)):
Inverse= higher levels of exposure associated with adverse health outcomes (score decreased)
Positive= higher levels of exposure associated with adverse health outcome (score increased)
Null = No association observed, p> 0.1
MDI = Mental development index
PDI = Psychomotor development index
Not tested = study did not measure the outcome at the age listed
N.I. = no information available
Collected, analysis not available = biomarker and outcome measured but associations never publicly released

Table 2. Urinary biomarkers of pesticide exposure.

Pesticide	Dimethyl-phosphate	Dimethylthio-phosphate	Dimethyldithio-phosphate	Diethyl-phosphate	Diethylthio-phosphate	Diethyldithio-Phosphate
Azinphos methyl	X	X	X			
Chlorethoxyphos				X	X	
Chlorpyrifos				X	X	
Chlorpyrifos methyl	X	X				
Coumaphos				X	X	
Dichlorvos	X					
Diazinon				X	X	
Dicrotophos	X					
Dimethoate	X		X			
Disulfoton				X	X	X
Ehtion				X	X	X
Fenitrothion	X	X				
Fenthion	X	X				
Isazaphos-methyl	X	X				
Malathion	X	X	X			
Methidation	X	X	X			
Methyl parathion	X	X				
Naled	X					
Oxydemeton-methyl	X	X				
Parathion				X	X	
Phorate				X	X	X
Phosmet	X	X	X			
Pirimiphos-methyl	X	X				
Sulfotepp				X	X	
Temephos	X	X				
Terbufos				X	X	X
Tetrachlorviphos	X					
Trichlorfon	X					

The table shows the six urinary metabolites and the parent organophosphate insecticides responsible for these metabolites. DAPs measures all six of these metabolites, only two of which are associated with chlorpyrifos. (CDC Fourth National Report on Human Exposure to Environmental Chemical, 2009)

REFERENCES

Columbia Study

Lovasi, G. S., Quinn, J. W., Rauh, V. A., Perera, F. P., Andrews, H. F., et al. (2011). Chlorpyrifos Exposure and Urban Residential Environment Characteristics as Determinants of Early Childhood Neurodevelopment. Am J Public Health 101:63-70.

Rauh V., Arunajadai S., Horton M., Perera F., Hoepner L., Barr D.B., Whyatt R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ Health Perspect 119:1196-201.

Rauh V.A., Garfinkel R., Perera F.P., Andrews H.F., Hoepner L., Barr D.B., Whitehead R., Tang D., Whyatt R.W. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics 118:e1845-59.

Whyatt R.M., Rauh V., Barr D.B., Camann D.E., Andrews H.F., Garfinkel R., Hoepner L.A., Diaz D., Dietrich J., Reyes A., Tang D., Kinney P.L., Perera F.P. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. Environ Health Perspect 112:1125-32.

Mt. Sinai

Berkowitz G., Wetmur J.G., Birman-Deych E., Obel J., Lapinski R.H., Godbold J., Holzman I.R., Wolff M. (2004). In Utero Pesticide Exposure, Maternal Paraoxonase Activity, and Head Circumference. Environ Health Persp 112:388-391.

Engel S.M., Wetmur J., Chen J., Zhu C., Barr D.B., Canfield R.L., Wolff M.S. (2011). Prenatal exposure to organophosphates, paraoxonase 1, and cognitive development in childhood. Environ Health Perspect 119:1182-8.

Berkeley (CHAMACOS)

Bouchard M.F., Chevrier J., Harley K.G., Kogut K., Vedar M., Calderon N., Trujillo C., Johnson C., Bradman A., Barr D.B., Eskenazi B. (2011). Prenatal Exposure to Organophosphate Pesticides and IQ in 7-Year Old Children. Environ Health Perspect 119:1189-1195.

Eskenazi B., Harley K., Bradman A., Weltzien E., Jewell N.P., Barr D.B., Furlong C.E., Holland N.T. (2004). Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. Environ Health Perspect 112:1116-24.

Eskenazi B., Marks A.R., Bradman A., Harley K., Barr D.B., Johnson C., Morga N., Jewell N.P. (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. Environ Health Perspect 115:792-8.

Huen, K.;Bradman, A.;Harley, K.;Yousefi, P.;Boyd Barr, D., et al. (2012). Organophosphate Pesticide Levels in Blood and Urine of Women and Newborns Living in an Agricultural Community. *Environ Res* 117:8-16.

Marks A.R., Harley K., Bradman A., Kogut K., Barr D.B., Johnson C., Calderon N., Eskenazi B. (2010). Organophosphate pesticide exposure and attention in young Mexican-American children: the CHAMACOS study. Environ Health Perspect 118:1768-74.

Other

Altman D.G., Bland J.M (2011). How to obtain a confidence interval from a P value. *Brit Med J* 342:d2090.

Barr D.B., Ananth C.V., Yan X., Lashley S., Smulian J.C., Ledoux T.A., Hore P., Robson M.G. (2010). Pesticide concentrations in maternal and umbilical cord sera and their relation to birth outcomes in a population of pregnant women and newborns in New Jersey. Sci Total Environ 408:790-5.

Burns, C. J.;McIntosh, L. J.;Mink, P. J.;Jurek, A. M.;Li, A. A. (2013). Pesticide Exposure and Neurodevelopmental Outcomes: Review of the Epidemiologic and Animal Studies. *J Toxicol Environ Health B Crit Rev* 16:127-283.

Fortenberry, G. Z.;Meeker, J. D.;Sanchez, B. N.;Barr, D. B.;Panuwet, P., et al. (2013). Urinary 3,5,6-Trichloro-2-Pyridinol (Tcpy) in Pregnant Women from Mexico City: Distribution, Temporal Variability, and Relationship with Child Attention and Hyperactivity. *Int J Hyg Environ Health* (in press).

Li, A. A.;Lowe, K. A.;McIntosh, L. J.;Mink, P. J. (2012). Evaluation of Epidemiology and Animal Data for Risk Assessment: Chlorpyrifos Developmental Neurobehavioral Outcomes. *J Toxicol Environ Health B* 15:109-184.

Mink, P. J.;Kimmel, C. A.;Li, A. A. (2012). Potential Effects of Chlorpyrifos on Fetal Growth Outcomes: Implications for Risk Assessment. *J Toxicol Environ Health B* 15:281-316.

Prueitt, R. L.;Goodman, J. E.;Bailey, L. A.;Rhomberg, L. R. (2011). Hypothesis-Based Weight-of-Evidence Evaluation of the Neurodevelopmental Effects of Chlorpyrifos. *Crit Rev Toxicol* 41:822-903

US EPA, FIFRA Scientific Advisory Panel. (2012). "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Chlorpyrifos Health Effects: Minutes of the FIFRA Science Advisory Panel Meeting held on April 10-12, 2012." SAP Minutes No. 2012-04. 108p.

US EPA, (2012). "Draft issue paper: Scientific issues concerning health effects of chlorpyrifos " April 10 – 13, 2012.

Wickerham, E. L.;Lozoff, B.;Shao, J.;Kaciroti, N.;Xia, Y., et al. (2012). Reduced Birth Weight in Relation to Pesticide Mixtures Detected in Cord Blood of Full-Term Infants. *Environ Int* 47:80-85.

Yan, X., Lashley, S., Smulian, J. C., Ananth, C. V., Barr, D. B., Ledoux, T. A., Hore, P., Robson, M. G. (2009). Pesticide Concentrations in Matrices Collected in the Perinatal Period in a Population of Pregnant Women and Newborns in New Jersey, USA, *Human Ecolog Risk Asses* 15:948 – 967.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF CHEMICAL SAFETY
AND POLLUTION PREVENTION

JAN 25 2013

Aaron Colangelo, Esq.
Senior Attorney
Natural Resources Defense Council
1152 15th Street NW, Suite 300
Washington, D.C. 20005

Margaret Reeves, Ph.D.
Senior Scientist/Program Coordinator (Environmental Health and Workers' Rights)
Pesticide Action Network North America
49 Powell Street, Suite 500
San Francisco, CA 94102

Re: Chlorpyrifos petition dated September 12, 2007; January 2013 Response

Dear Mr. Colangelo and Dr. Reeves:

I am writing to further update you on the U.S. Environmental Protection Agency's (EPA) efforts to respond to the Natural Resources Defense Council (NRDC) and Pesticide Action Network North America (PANNA) jointly submitted September 12, 2007[1], petition and our related efforts to complete the registration review of chlorpyrifos. In my letter to you of December 18, 2012[2], I provided you with an update on our efforts to implement label changes to put in place additional limitations to reduce primary spray drift from chlorpyrifos. I can report that EPA has now approved those changes for all 41 chlorpyrifos agricultural products subject to these use limitations.

As we also noted in December, while we have made significant progress in completing work on the four petition issues that EPA did not address in its July 16, 2012[3], partial response to your petition, we were not able to provide you with a complete response in December, as we previously believed we could. However, we committed to providing you with a response this month that further addresses the petition and outlined the approach we are taking for completing our response. Accordingly, this response will address what EPA has done and will do to address each of the following four outstanding claims that: (1) EPA failed to incorporate inhalation routes of exposure from pesticide volatilization; (2) EPA failed to incorporate into its risk

[1] Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2007-1005-0005
[2] Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2007-1005-0096
[3] Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2007-1005-0095.

assessment, in a quantitative manner, data indicating that long-lasting effects result from early life exposure to chlorpyrifos in children; (3) EPA disregarded data demonstrating that there is no evidence of a safe level of exposure during pre-birth and early life stages; and (4) EPA failed to cite or quantitatively incorporate studies and clinical reports suggesting potential adverse effects below 10% cholinesterase inhibition.

As I indicated in the December response, EPA has been working to complete an assessment that will evaluate the potential risks of volatilization from chlorpyrifos applications. In early February 2013, we will publish a notice in the Federal Register announcing the availability of this preliminary assessment for public comment. This assessment represents a significant advancement in the evaluation of pesticide risks, as it will be the first probabilistic assessment of the risks posed by the post-application volatilization of a semi-volatile pesticide. Our approach builds upon the methodology we previously employed for volatile pesticides in the recent fumigant pesticide risk assessments[4] to assess bystander inhalation exposure from volatilization. In addition, it is consistent with the recommendations from the December 2009 Federal Insecticide, Fungicide, and Rodenticide Act Scientific Advisory Panel (SAP)[5] meeting on the scientific issues associated with field volatilization of conventional (semi-volatile) pesticides. The content of the preliminary volatilization assessment is further informed by Dow AgroSciences' recently submitted chlorpyrifos field volatility study[6] coupled with existing volatility data found in the open literature, and EPA modeling tools.

This assessment will supplement the July 2011 Preliminary Human Health Risk Assessment[7] (HHRA) and evaluates bystander exposure from chlorpyrifos and chlorpyrifos-oxon emitted from treated fields. Although the volatilization of chlorpyrifos was addressed in the preliminary HHRA, that analysis involved only a deterministic assessment based on limited monitoring data that did not attempt to evaluate a range of field conditions and, therefore, had correspondingly limited utility in a regulatory setting. Given the groundbreaking nature of the new assessment and its potential for use in guiding additional risk mitigation, EPA believes it is critical to involve the public in the development of this assessment before it is finalized. Further, EPA is examining other semi-volatile pesticides to determine if bystander volatilization assessments are needed. Any comments received on this assessment will serve to inform those assessments as well. Accordingly, EPA will begin taking public comment on the draft version of the assessment in February 2013, after publication of the Federal Register notice announcing its availability in docket number EPA-HQ-OPP-2008-0850.

Following completion of the public comment period and EPA's subsequent evaluation of the comments, EPA will determine whether additional regulatory action is necessary to address

---

[4] The assessments can be found in the dockets for each fumigant. Four of which are provided here chloropicrin - EPA-HQ-OPP-2007-0350; dazomet - EPA-HQ-OPP-2005-0128; metam sodium/potassium - EPA-HQ-OPP-2005-0125; and methyl bromide - EPA-HQ-OPP-2005-0123

[5] U.S. EPA 2009. FIFRA Science Advisory Panel Meeting Minutes - Scientific Issues Associated with Field Volatilization of Conventional Pesticides. Available at http://www.epa.gov/scipoly/sap/meetings/2009/december/120309meetingminutes.pdf

[6] Rotondaro, A. and Havens, P. (2012). Direct Flux Measurement of Chlorpyrifos and Chlorpyrifos-Oxon Emissions Following Applications of Lorsban Advanced Insecticide to Alfalfa; Sponsor: Dow AgroSciences LLC, 9330 Zionsville Road Indianapolis, IN 46268-1054. EPA MRID 48883201.

[7] Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2008-0850-0025.

these risks and, if so, whether the nature of that risk supports the need to take action in advance of our completion of the final broader HHRA, currently scheduled for December 2013.

Regarding the remaining three petition issues addressing chlorpyrifos toxicity identified above, as we have indicated previously, the analysis is complicated and multifaceted because it involves many lines of scientific evidence, including many recently conducted studies and peer review evaluations and recommendations. That work includes consideration of: *in vivo* and *in vitro* experimental toxicology studies that evaluate neurodevelopmental effects in laboratory animals, adverse outcome pathway framework analyses, exposure, the results of multiple human epidemiology studies, and biomonitoring data. Notwithstanding the complexity of this analysis, it was our hope to provide you with a written response last December that included our scientific conclusions on these issues. As you know, we convened a FIFRA SAP meeting in April 2012[8] to inform our work in generating a weight-of-evidence evaluation integrating the epidemiologic data with the experimental toxicology studies for the neurodevelopmental outcomes and acetylcholinesterase (AChE) inhibition. At the time EPA provided its partial petition response to you in July 2012, EPA had just received the written SAP report from the April meeting. EPA therefore had not had time to begin pursuing the SAP's recommendation when EPA provided its response to you and to the 9th Circuit in our ongoing litigation over this matter.

Thus far, EPA has not encountered epidemiological data of sufficient quality to support quantitative risk assessment of conventional pesticide chemicals. Before EPA decides how to use the epidemiological data on chlorpyrifos, we believe it is critical to attempt to resolve questions about these studies regarding the extent of the cohort members' exposures to chlorpyrifos, as well as the impact of exposure to other compounds capable of causing or contributing to the observed neurological outcomes. We acknowledge the lengthy conduct of our assessment, including multiple SAP reviews, but we believe the deliberate and considered approach we are taking is the most scientifically defensible method for re-evaluating our current approach to assessing risks from chlorpyrifos and other organophosphorous pesticides generally, and, specifically, for evaluating the strengths and weaknesses of the epidemiological data.

The July 2012 SAP report is in accord with EPA's assessment that the Agency should attempt to resolve certain key questions about the epidemiological data. Specifically, the SAP recommended that EPA pursue a number of possible approaches for attempting to resolve whether the neurological outcomes observed in the studies occurred in the absence of AChE inhibition – the effect EPA's current regulatory approach is designed to preclude. Further, given that the women and children studied in the Columbia University-sponsored epidemiology study[9] were exposed to multiple chemicals (including other pesticides, polycyclic aromatic hydrocarbons and lead), the SAP cautioned the agency about attributing the outcomes to a single chemical based on the current analysis conducted by Columbia University researchers. These statements by the SAP lead the agency to believe that we need to further explore the extent to

---

[8] Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2012-0040-0029

[9] Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D. B., & Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ Health Perspect, 119(8), 1196-1201. doi: 10.1289/ehp.1003160; Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., Barr, D. B., Whyatt, R. W. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics, 118(6), e1845-1859. doi: 10.1542/peds.2006-0338.

which the observed neurological outcomes were influenced by exposure to these other chemicals.

Following receipt of the report EPA began conducting a number of analyses to address these recommendations. As I indicated in our December response, we are making progress in conducting a dose-reconstruction analysis of potential exposures to the women and children studied in the Columbia University-sponsored epidemiology study[10] in order to assess the degree to which the individuals in the cohort may or may not have been exposed to chlorpyrifos levels high enough to cause AChE inhibition. In addition to this assessment, to address the SAP recommendations EPA also intends in the coming months to complete an evaluation of cohort exposures to other chemicals. In order to complete both the dose reconstruction and analyses on other chemical exposures, however, we will need to analyze the original data ("raw data") from the Columbia University study to better understand the exposure to chlorpyrifos and other chemicals. To date, the study authors have declined our request to provide that information to us, but we are continuing to discuss our need for evaluating these data with the study authors and we are hopeful that a resolution can be reached.

In addition to further analysis of the exposures in the Columbia study, EPA has also followed up on a recommendation that was brought up in the SAP's oral deliberations regarding the administration and interpretation of diagnostic and analytic tools used to assess neuro and motor development in children like those used in the Columbia study. The SAP noted that it lacked expertise in evaluating these aspects of the data. Because this expertise is relevant in assessing the potential for effects from exposures to other chemicals, between August and October 2012, we obtained additional peer review from scientists within the federal government who have expertise in this field. EPA will include consideration of the results of this peer review when it completes its assessment, as further discussed below.

Finally, as our previous response indicated, last fall, the Columbia University researchers published a new epidemiology study[11] describing the results of magnetic resonance imaging on a subset of children in the cohort. We solicited comments between August 2012 and October 2012, from scientists within the federal government who have expertise in this scientific area and are currently evaluating this input to determine the extent to which this information informs the earlier Columbia University study results.

In light of our ongoing work described above, we are not in a position to provide you with our conclusions on the three remaining toxicology issues in the petition at this time, and it is difficult to provide a precise time frame for the completion of that assessment. It is our hope that we can maintain our current schedule to complete the full chlorpyrifos HHRA by the end of 2013

---

[10] Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D. B., & Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ Health Perspect, 119(8), 1196-1201. doi: 10.1289/ehp.1003160; Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., Barr, D. B., Whyatt, R. W. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics, 118(6), e1845-1859. doi: 10.1542/peds.2006-0338.

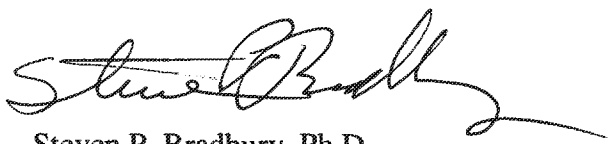
[11] Rauh VA, Perera FP, Horton MK, Whyatt RM, Bansal R, Hao X, Liu J, Barr DB, Slotkin TA, Peterson BS. Brain anomalies in children exposed prenatally to a common organophosphate pesticide. Proc Natl Acad Sci U S A. 2012 May 15;109(20):7871-6. doi: 10.1073/pnas.1203396109. Epub 2012 Apr 30. PubMed PMID: 22547821; PubMed Central PMCID: PMC3356641.

and respond to the remaining claims in your petition on the same time frame. As we previously explained to you, that schedule would result in our initiating any necessary regulatory action in early 2014. Given the complexity of the assessment, and in particular, the complications we are having in obtaining potentially important research data from the Columbia University study authors, I do have some concern about our ability to meet that time frame, but we will continue to work to meet that goal and will update you if our plans must change.

With that said, we have made significant progress in addressing the volatilization portion of your inhalation claim as will be evident with the release of the preliminary chlorpyrifos volatilization assessment in February. As noted, if, following review of the public comments, EPA determines that the risk posed from chlorpyrifos volatilization merits regulatory action in advance of the completion of the HHRA, we will initiate that action without first completing the entire HHRA.

Finally, I wish to reiterate that for efficiency purposes EPA does not intend to proceed with issuing a denial order of the six petition issues (the spray drift portion of your inhalation claim was granted) that we rejected in July 2012 until after we complete our review of all remaining issues. It has been our understanding that this approach is preferable to you as well. However, as previously indicated, if you wish to begin the objections process for the six denied claims and notify EPA in writing, we will publish a formal denial order for those claims, triggering your right to file objections under FFDCA section 408(g)(2).

Sincerely



Steven P. Bradbury, Ph.D.
Director, Office of Pesticide Programs

Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, 2015; 45(7): 531–641
 DOI: 10.3109/10408444.2015.1043976

informa
healthcare

REVIEW ARTICLE

A review of epidemiologic studies of low-level exposures to organophosphorus insecticides in non-occupational populations

Richard Reiss[1], Ellen T. Chang[2,3], Rudy J. Richardson[4,5], and Michael Goodman[6]

[1]*Exponent, Alexandria, VA, USA,* [2]*Exponent, Menlo Park, CA, USA,* [3]*Division of Epidemiology, Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA,* [4]*Department of, Environmental Health Sciences,* [5]*Department of Neurology, University of Michigan, Ann Arbor, MI, USA and* [6]*Department of Epidemiology, Emory University, Rollins School of Public Health, Atlanta, GA, USA*

Abstract

This paper systematically reviews epidemiologic studies related to low-level non-occupational exposures to organophosphorus (OP) insecticides. Many of the studies evaluate levels of maternal OP metabolites and subsequent health outcomes in offspring. The studies focused primarily on birth outcomes (e.g., infant body weight or head circumference) and neurodevelopmental (e.g., mental and psychomotor) testing results. The evidence from these studies was reviewed under the Bradford Hill guidelines. Most of the studies assessing exposure based on urinary levels of OP insecticide metabolites used only one or two measurements during pregnancy. The potential for exposure misclassification with this method is largely due to (1) preformed metabolites that are ingested with food, (2) the short elimination half-life of OP insecticides, and (3) lack of specificity to particular OP insecticides for many of the metabolites. For birth outcomes, the majority of reported results are not statistically significant, and the associations are inconsistent within and across studies. There is more within-study consistency for some of the neurodevelopmental testing results, although few associations were examined across several studies. These associations are generally weak, have been replicated only to a limited extent, and require further confirmation before they can be considered established. The OP insecticide levels measured in the epidemiologic studies are too low to cause biologically meaningful acetylcholinesterase inhibition, the most widely used metric for OP insecticide toxicity. Overall, the available evidence does not establish that low-level exposures to OP insecticides cause adverse birth outcomes or neurodevelopmental problems in humans.

Keywords

Bradford Hill, epidemiology, insecticides, organophosphorus, pesticides

History

Received 13 November 2014
Revised 14 April 2015
Accepted 19 April 2015
Published online 7 July 2015

Table of Contents

Address for correspondence: Richard Reiss, Exponent, 1800 Diagonal Road, Suite 500, Alexandria, VA 22310, USA. Tel: + (571) 227-7228. E-mail: rreiss@exponent.com

Introduction

Organophosphorus (OP) insecticides, or their oxon metabolites, persistently inactivate acetylcholinesterase (AChE), an enzyme involved in neurotransmission in insects as well as humans and other animals. OP insecticides are used widely around the world. Most studies of the adverse human health effects of exposure to OP insecticides have focused on occupational or other high-dose exposures, including acute poisoning. Acute clinical effects result from AChE inhibition at synapses in the central nervous system, autonomic nervous system, and neuromuscular junction (Eddleston et al. 2008).

Over the last decade, a number of epidemiologic studies have been published that evaluate the potential health effects of OP insecticides in populations with little or no occupational exposure. These epidemiologic studies have most frequently evaluated birth outcomes, such as infant body weight and head circumference, or results of neurodevelopmental tests that measure mental and psychomotor function. The primary exposure pathways for subjects in these studies likely include diet, residential use, and in some cases, proximity to agricultural operations. Exposures in these studies are estimated primarily by measuring OP insecticide biomarkers or degradation products (referred to in this article as "OP metabolites") in urine and blood. In several study populations, such markers have been measured at levels that are sufficiently low to indicate that exposure to OP insecticides originates predominantly from dietary sources (Berman et al. 2013, Lu et al. 2008).

Risk assessments in Europe and the United States have concluded that dietary exposure to OP insecticides appears generally to be safe (Boon et al. 2008, Claeys et al. 2008, Jensen et al. 2003, Jensen et al. 2009, Nougadere et al. 2012). Nevertheless, several recent epidemiologic studies that measured OP metabolites in blood or urine suggest associations between low-dose exposure to OP insecticides and adverse human health effects. Most of these studies have focused on OP insecticide metabolite levels *in utero*, which is believed to be the critical exposure period for human neurological development (Rice and Barone 2000) and is, by definition, the only relevant exposure period for birth outcomes. Given the widespread use of OP insecticides and consumption of OP-treated foods, understanding the potential human health impact of low-dose exposure to OP insecticides is important from a public health and regulatory standpoint. We undertook this systematic review of epidemiologic studies of low-level OP metabolites to evaluate the existing evidence on associations with adverse human health outcomes. A few previous papers have reviewed the epidemiologic literature specific to chlorpyrifos for neurobehavioral outcomes (Li et al. 2012) and fetal growth outcomes (Mink et al. 2012), and found no compelling evidence of effects. Burns et al. (2013) reviewed animal toxicology and epidemiologic data for neurodevelopmental outcomes and all classes of pesticides. The researchers found that the epidemiologic literature did not support causal effects for pesticides, and that effects found in toxicology studies were generally seen at doses similar to or higher than points of departure used in regulatory risk assessments. This review is the first to address potential effects of all OP insecticides from epidemiologic studies with low-level exposures.

To evaluate the scientific evidence for a conclusion regarding causality, we used the Bradford Hill guidelines, including strength of association, consistency, temporality, biological gradient, plausibility, coherence with toxicological evidence, specificity, experiment, and analogy (Hill 1965). The manuscript also includes a detailed evaluation of the validity of the urinary biomarkers used in the epidemiologic studies, and reviews the plausibility of the associations by comparing animal and limited human toxicology data with the OP insecticide levels observed in the epidemiologic studies. Potential confounding and bias are also evaluated. The data are then assembled to assess overall evidence for and against a causal relationship between low-level exposure to OP insecticides and adverse birth outcomes or neurodevelopmental problems in humans.

Scope of review

To identify the relevant studies on low-level OP metabolites and human health outcomes, we used PubMed to search MEDLINE using keywords and keyword roots, including *organophosph**, specific metabolites (e.g., *dialkylphosphate** or *dialkyl phosphate*), specific OP insecticides (e.g., *chlorpyrifos*, *diazinon*, *malathion*, *parathion*, or *phosmet*), and various age groups (e.g., *child**, *infan**, *toddler**, *birth**, *men*, *women*, or *adult**). Based on a review of titles and abstracts, we excluded more than 1500 articles that presented animal and *in vitro* studies, biomonitoring studies, and other non-epidemiologic studies, including case reports, commentaries, and reviews (some of which were examined to identify references missed by the electronic search). After reviewing full-text articles, we further excluded 40 studies of occupational or para-occupational (i.e., take-home) exposure to OP insecticides, exposure by poisoning, exposure by pediculosis treatment, exposure by aerial residential or illegal indoor residential spraying, exposure to pesticides or insecticides not specific to OP compounds, and paraoxonase 1 (*PON1*) genotype or PON1 enzyme activity without specific evaluation of OP insecticide exposure. We further excluded 31 studies that estimated OP exposure based on self-reported or geographic data, and those that estimated associations with health categories that were evaluated in fewer than three independent studies, thereby providing an insufficient basis for a weight-of-evidence evaluation. Based on this last consideration, the two endpoint categories of interest in this review are birth outcomes and results of neurodevelopmental testing. We ultimately included 31 epidemiologic studies—11 studies of birth outcomes and 20 studies of neurodevelopmental outcomes—in this review.

Study characteristics—including study name, location, design, description and number of subjects, follow-up time, exposure assessment methods, outcome assessment methods, point and interval estimates of association between specific exposures and outcomes of interest, and adjustment factors—were abstracted from each relevant study, and independently checked by another reviewer for accuracy. Individual studies were evaluated with respect to strength of study design, exposure and outcome assessment, potential for confounding

and bias, role of random error or chance, and interpretation of results. To evaluate the overall weight of epidemiologic evidence, we used the framework of the Bradford Hill guidelines (Hill 1965). The Bradford Hill guidelines are one of the most common and established methods of assessing evidence for a causal relationship between an exposure and a disease (Gordis 2013). We assessed the study results relative to each of the Bradford Hill guidelines, separately for birth outcomes and neurodevelopment. These aspects were used as considerations, but not as strict criteria in a checklist fashion, to guide our evaluation of causality.

Overarching issues

Before proceeding to a review of the individual studies, three overarching issues need to be discussed. First, most studies use urinary levels of OP insecticide metabolites to classify exposures. Therefore, we discuss the validity of exposure assessment using these urinary biomarkers. Second, the OP insecticide exposure levels of the study subjects are generally lower than those previously identified as harmful. Therefore, we briefly review the extensive animal toxicology and limited human toxicology data to evaluate the plausibility of the associations observed in epidemiologic studies. Third, in any epidemiologic study, confounding and bias should be considered as potential explanations for an observed result (Gordis 2013).

Biomarker validity

Most epidemiologic studies that use biomarkers of OP insecticide exposure rely on urinary measurements of OP metabolites. The metabolites include six dialkylphosphates (DAPs): dimethylphosphate (DMP), dimethylthiophosphate (DMTP), dimethyldithiophosphate (DMDTP), diethylphosphate (DEP), diethylthiophosphate (DETP), and diethyldithiophosphate (DEDTP). The first three are commonly grouped as DMPs, and the latter three are commonly grouped as DEPs. In a few studies, metabolites of specific OP insecticides (usually chlorpyrifos or malathion) were measured. Use of these metabolites as exposure biomarkers has the potential for exposure misclassification, for the following reasons:

- OP metabolites formed directly on or in food are well absorbed orally and cannot be distinguished from those formed following absorption.
- Rapid metabolism of OP insecticides and their metabolites results in high intra-individual variation in levels, such that single samples may not reflect past or long-term average exposure.
- DAP metabolites are not specific to individual OP insecticides, and there is a vast range of toxicity across different compounds.
- Variability in exposure measurements across studies diminishes the comparability of results.

Direct formation of OP metabolites in food

DAPs are products of OP hydrolysis. The metabolism of OP insecticides in plants and humans is similar. The DAPs detected in human urine may therefore have been ingested with food or formed in the body following absorption of OP insecticides (e.g., Zhang et al. 2008). Compared with their parent compounds, DAPs are virtually non-toxic (Chen et al. 2013). Some studies used chemical-specific urinary metabolites, including malathion dicarboxylic acid (MDA, a metabolite of malathion) and 3,5,6-trichloro-2-pyridinol (TCP, a metabolite of chlorpyrifos), as OP insecticide exposure biomarkers. These chemical-specific metabolites also form directly in food (Morgan et al. 2011, Chen et al. 2012) and are also relatively non-toxic compared with their parent compounds or active metabolites (Chen et al. 2012, Eaton et al. 2008)[1].

Several studies have demonstrated that most dietary exposures are actually to the OP metabolites and not to the parent compounds. Zhang et al. (2008) measured OP and DAP levels on 153 produce samples known to be contaminated with OP insecticides. The mean concentrations of OP insecticides and DAP residues were 1.2 and 2.0 nmol/g, respectively. On a molar basis, more than 60% of the total residues were DAPs. In addition, 60% of the samples contained higher DAP than OP residues. The mole fraction of DAPs across the samples varied widely, ranging from 0.02 to 0.99. Zhang et al. (2008) found that the mole ratio of DAPs to parent OP insecticides was both produce-specific and chemical-specific, with higher ratios for diazinon, phosmet, chlorpyrifos, azinphos-methyl, and malathion. When measured on strawberries, the ratio of DAPs to parent insecticide (malathion) increased with time since application, indicating continuous transformation. The mole ratio of DAPs to malathion was 1.4 one day after application, and increased to 8.7 after 9 days.

Morgan et al. (2011) measured chlorpyrifos and TCP levels in food from homes and daycare centers of 127 Ohio preschool children. The mean chlorpyrifos residues were 0.4 ng/g in homes ($n = 125$) and 0.2 ng/g in daycare centers ($n = 29$). The mean TCP residues were 2.6 ng/g in homes ($n = 127$) and 2.8 ng/g in daycare centers ($n = 29$). Thus, the TCP residues were significantly higher than the chlorpyrifos residues. Moreover, the Pearson correlation coefficient for dietary chlorpyrifos and excreted urinary TCP was only 0.30, meaning that dietary chlorpyrifos exposure explained only about 9% of the variability in excreted urinary TCP.

Chen et al. (2012) measured malathion and its transformation products, including the DAPs, MDA, and malathion monocarboxylic acid (MMA), in 157 produce samples. The samples had been confirmed previously to contain detectable malathion, but no detectable levels of other OP insecticides. The mean malathion residue was 0.60 nmol/g, and the mean preformed metabolite residue was 3.29 nmol/g. The mole fraction of preformed metabolites (DAP + MMA + MDA) ranged from 0.41 to 1.00. The mole ratio of total metabolites to malathion parent ranged from 0.70 to 333.

In summary, by demonstrating that most of the DAPs, MDA, and TCP are formed on food items, these studies indicate that the metabolite concentration measured in urine may be due to direct exposure to these relatively non-toxic compounds, rather than to the parent OP insecticide. The substantial variability in the metabolite-to-parent ratio reduces the value of excreted metabolites as markers of OP insecticide exposure.

[1]To be precise, when formed in the environment, the DAPs are not "metabolites" formed by enzymatic transformations, but rather are degradation products formed by hydrolysis or photolysis. However, we use the term "DAP metabolites" in the paper for brevity.

Rapid metabolism of OP insecticides

The epidemiologic studies of prenatal OP exposure typically include either one or two urinary measurements of OP metabolites that are intended to represent the exposure of the mother during pregnancy. However, many OP insecticides are metabolized relatively rapidly. Most OP insecticides are typically excreted within 24–48 h (World Health Organization [WHO] 1996). Some human exposure data suggest even faster rates of metabolism for particular OP insecticides. For example, Garfitt et al. (2002) reported that a single oral dose of diazinon has a urinary elimination half-life of 2 h. In a similar study, Bouchard et al. (2003) estimated a 4-h half-life for malathion.

Given the rapid elimination of OP insecticides, any spot measurement will reflect only recent exposure. If the relevant exposure period of interest is an average over pregnancy, a single measurement may be inadequate. There is no biological basis to specify a particular exposure period during pregnancy as especially relevant for neonatal or childhood outcomes examined in this review. However, if the exposure period of interest is a short time window during pregnancy, then a spot measurement taken outside that window may not be etiologically relevant.

DAP metabolites are non-specific

Multiple OP insecticides are metabolized into each of the six DAPs (Duggan et al. 2003, Sudakin and Stone 2011). Some OP insecticides (e.g., malathion and disulfoton) are converted to as many as three different DAPs, whereas others (e.g., dichlorvos and tetrachlorvinphos) metabolize to only a single DAP. Moreover, acephate and methamidophos do not metabolize to DAPs at all (Solecki 2002).

There are substantial differences in toxicity across the OP insecticides. The U.S. Environmental Protection Agency (EPA) estimated chronic exposure benchmark doses using 10% brain AChE inhibition threshold (BMD_{10}) for all registered OP insecticides. AChE inhibition is the widely recognized mechanism of action for OP toxicity (Mileson et al. 1998). The BMD_{10} values in the EPA assessment, based on rat laboratory studies, ranged from 0.04 milligrams per kilogram body weight per day (mg/kg/day) for dicrotophos to 313.9 mg/kg/day for malathion (USEPA 2002), a nearly 8000-fold difference. Even among the most widely used OP insecticides, the toxicity varies over orders of magnitude (see next section). Such large differences in toxicity across OP insecticides, combined with the lack of specificity for DAPs, significantly limit the ability of DAP urinary levels to provide an informative measure of toxic exposure.

Intra-individual variability in urinary DAP levels

Studies with repeated measures of urinary DAP concentrations offer useful information on intra-individual variability. Bradman et al. (2013) found that spot DAP measurements in children changed up to two orders of magnitude over a week or even within a day. In 24-h urine samples, the DAP levels differed by as much as an order of magnitude for samples collected three days apart.

A number of researchers have reported that within-child variability in DAP levels is higher than between-child variability (Griffith et al. 2011, Sexton and Ryan 2012, Bradman et al. 2013, Attfield et al. 2014). For example, within-child variability in one study was 2–11 times greater than that observed across the study population (Attfield et al. 2014). Griffith et al. (2011) found similar results for children living in an agricultural community in central Washington State. Sexton and Ryan (2012) measured the intraclass correlation coefficient for urinary DAP among elementary school children in Minneapolis, and observed "only modest correlations" in siblings from the same household.

Because the associations estimated in the epidemiologic studies are based on one, or at most two, DAP measurements, a higher level of intra- than inter-individual variability can lead to considerable exposure misclassification. Attfield et al. (2014) illustrated this problem by assigning subjects with multiple available OP metabolite measures to four exposure categories based on the mean values of 1–4 randomly selected samples. If the metric under study is reliable, the grand means of the four resulting exposure categories are expected to increase monotonically. In this study, however, the resulting grand means were monotonic only 14–15% of the time for MDA and 19–32% for TCPy, when the exposure assessment was based on only one sample per subject. When two samples were used, the resulting grand means for MDA and TCPy were monotonic 31–32% and 34–41% of the time, respectively.

Potential impact of exposure misclassification

It is important to consider the potential impact of misclassification of OP insecticide exposure on the results of epidemiologic studies. It is often said that if exposure misclassification is non-differential (i.e., independent of health status), bias is expected to produce an attenuated measure of association (Cantor et al. 1992). However, it is plausible that short-term variability in dietary patterns and other influences on OP insecticide exposure differ by health status. For example, diet is associated with birth outcomes and neurodevelopment (Abu-Saad and Fraser 2010, Millichap and Yee 2012, Smithers et al. 2013), and changes in diet are commonly triggered by health status. Consequently, if diseased individuals altered their dietary habits more frequently than non-diseased individuals, then the degree of exposure misclassification would differ by health status, leading to an unknown degree or direction of bias. Even if exposure misclassification is non-differential by health outcome, it does not necessarily result in a predictable direction of bias. Additional conditions, such as independence of classification errors, must be met for non-differential misclassification of a binary exposure to result in bias toward the null, and even then the tendency applies only to the expectation of the estimated association, not to the value of the estimate from any single study (Jurek et al. 2008, Jurek et al. 2005). Moreover, for exposures with multiple levels, non-differential misclassification results in bias of unpredictable direction and magnitude (Sorahan and Gilthorpe 1994, Wacholder et al. 1995).

Dose–response

OP insecticides or their active metabolites inhibit the enzyme AChE, which breaks down the neurotransmitter acetylcholine. Neurotoxicity results from excessive accumulation of acetylcholine in cholinergic synapses. Thus, inhibition of

nervous system AChE is generally regarded as the primary toxic mode of action for OP insecticides (Mileson et al. 1998, U.S. EPA 2000). Accordingly, the U.S. EPA regulates OP insecticide safety by setting exposure levels to be sufficiently low that excessive AChE inhibition will not occur (U.S. EPA 2000). It is possible that developmental neurotoxicity may result from mechanisms other than AChE inhibition (Yang et al. 2011). However, the U.S. EPA requires developmental neurotoxicity studies for OP insecticides and has found that AChE inhibition is protective of developmental neurotoxicity effects. It is acknowledged that developmental neurotoxicity studies in animals may not be sensitive enough to detect all developmental neurotoxicity-related effects; research in this area continues.

In humans and other mammals, AChE exists in both the nervous system (brain, spinal cord, and peripheral plexuses and nerves) and the red blood cells (RBCs) with varying amounts in plasma in some species. Another type of cholinesterase, butyrylcholinesterase (BChE), is found in plasma and other tissues (Li et al. 2005).

Inhibition of blood cholinesterase, either in RBCs or plasma, is generally regarded as a marker of exposure, but not necessarily a toxic effect (U.S. EPA 2000). Nevertheless, because data on AChE activity in peripheral nervous system tissues may be lacking in animal studies and data on peripheral nervous system tissues and/or brain is usually lacking in humans, the EPA regards AChE inhibition in blood as a surrogate for peripheral nervous system AChE inhibition in animals and brain AChE inhibition in humans. Given that the relevant target for toxicity is nervous system AChE and extensive data are available on inhibition of brain AChE in rats and other non-human species, the focus of the analysis described below is on brain AChE inhibition.

It is useful to examine these data relative to the OP biomarker levels in non-occupational settings to determine the potential for brain AChE inhibition at the exposure levels found in the epidemiologic studies. As part of its risk assessments for registration review, the U.S. EPA has developed AChE dose–response models for brain AChE for OP insecticides used in the United States. The dose–response models are based on the benchmark dose for 10% inhibition (BMD_{10}) of brain AChE in animal studies. The BMD_{10} represents the dose that, on average across the animals, causes 10% AChE inhibition and is considered by the U.S. EPA to be a "response level close to the background cholinesterase" (U.S. EPA 2002). The dose–response models are based on an exponential decline of AChE activity with dose.

We reviewed the U.S. Department of Agriculture (USDA) Pesticide Data Program database to identify the OP insecticides most commonly detected in food. The latest data are from 2012 (USDA 2014). Four OP compounds—dimethoate, omethoate, malathion, and chlorpyrifos—account for nearly 80% of the 663 detections. In the U.S. EPA risk assessments (U.S. EPA 2005, 2009a, 2011), the lowest BMD_{10} values were 1.4 mg/kg (4.1 nmol/kg) for chlorpyrifos, 1.5 mg/kg (6.6 nmol/kg) for dimethoate, and 23.6 mg/kg (71.5 nmol/kg) for malathion. For omethoate, we used the BMD_{10} of 0.14 mg/kg (0.68 nmol/kg) based on a cholinesterase study conducted after the last U.S. EPA risk assessment (Reiss 2012). U.S. EPA used a slightly higher value of 0.18 mg/kg in its last dimethoate risk assessment based on earlier data (U.S. EPA 2005b). All of the BMD_{10} values are for exposure to rat pups on postnatal day 11 and are the lowest BMD_{10} estimates observed in pups, adults, and pregnant dams. The rat pups were exposed directly on postnatal day 11 and prenatally through exposure from the dam. The rats in these studies were generally well nourished, which may lead to uncertainty in applying the results to poorly nourished human populations.

It is useful to estimate exposures associated with DAP levels measured in the epidemiologic studies so that AChE inhibition associated with those DAP levels can be estimated. This can be roughly accomplished by back-calculating an exposure based on the DAP level and urine volume, acknowledging the uncertainties in the calculation. Curl et al. (2003) provides a simple equation to estimate the dosage associated with a urinary DAP measurement:

$$Dosage = \frac{DAP \times V \times MW}{BW}$$

where [DAP] is the total molar DAP concentration, V is the daily urine volume, MW is the molecular weight, and BW is the body weight. We assume a normal urine volume of 20 mL/kg/day (Gonzales and Bauer 1999). The urinary levels are corrected for DAPs formed on food items by assuming that 38% of the urinary DAP levels are from exposure to the pesticide, based on data from Zhang et al. (2008). Use of the above equation to estimate the dosage of OP insecticide associated with DAP measurements in the epidemiologic studies has important limitations. The DAPs originate from different OP compounds, but to apply the dose–response models, we need to assume that all DAPs originate from exposure to one OP insecticide. In addition, data on DAPs formed on food items are not available for all OP insecticides and commodities. The DAP measurements in the epidemiologic studies are typically spot samples, yet the equation estimates full-day exposures. Despite these limitations, the models provide a useful approximation to assess AChE inhibition for dosages corresponding to the urinary metabolite levels found in the epidemiologic studies.

Among participants in the 2000–2004 National Health and Nutrition Examination Survey (NHANES), the geometric mean of urinary DAP concentrations was 68 nmol/L, and the corresponding 75th percentile was 186 nmol/L (Bouchard et al. 2010). The NHANES data represent a sample of the general non-institutionalized U.S. population. For the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) epidemiologic study, Bradman et al. (2005) reported median and 90th percentile levels of 103 and 732 nmol/L for the first prenatal sample, 107 and 422 nmol/L for the second prenatal sample, and 227 and 1349 nmol/L for the postpartum sample, respectively. The most recent (2007–2008) NHANES data on urinary DAPs show that the 50th, 75th, 95th, and 99th percentiles across 2564 samples were 48, 155, 587, and 1406 nmol/L, respectively, assuming half the limit of detection for non-detects (CDC 2014). The 75th percentile of 155 nmol/L is somewhat lower than the 75th percentile reported by Bouchard et al. (2010) for the 2000–2004 NHANES data. The 98th percentile in the 2007–2008 NHANES data set is about 2000 nmol/L, which corresponds

to exposures of less than about 8–14 μg/kg/day, depending on the molecular weight of the OP compound.

Based on the dose–response models assuming that all exposures are from a single OP insecticide, at 2000 nmol/L, the estimated brain AChE inhibition was 0.002% for malathion and 0.001% for chlorpyrifos. While malathion has a higher BMD_{10}, the chlorpyrifos data were fit to a different dose–response model that has a low-dose shoulder, limiting inhibition at low doses. At higher doses, the models diverge, and malathion is estimated to cause less inhibition than chlorpyrifos. The estimated brain AChE inhibition at 2000 nmol/L is 0.03% for dimethoate and 0.2% for omethoate.

These low levels of brain AChE inhibition are highly unlikely to be clinically detectable, particularly considering the variety of factors that may affect AChE activity. For example, solanaceous glycoalkaloids found in potatoes cause AChE inhibition (Krasowski et al. 1997); so does huperzine, another natural product derived from club moss, which is used in the treatment of dementia (Ozarowski et al. 2013). The inhibition of AChE activity associated with huperzine is hypothesized to result in improvements in long-term memory (Ozarowski et al. 2013). Lefkowitz et al. (2007) evaluated baseline RBC AChE activity for 46 workers over an average of 20 years of employment. The mean coefficient of variance for RBC AChE was 3.9%. Ferioli and Maroni (2011) report inter-individual variations in RBC AChE of 10–18% and intra-individual variations of 3–7%. This baseline variance for individuals is higher than the estimated AChE inhibition at upper percentiles of the doses reported in the epidemiologic studies. Moreover, these data are for RBC AChE, which adds uncertainty, because RBC AChE activity serves as a surrogate measure of brain AChE function.

The estimates from the U.S. EPA dose–response models are for the mean response in rats and do not account for intra-individual variability or the potential for increased sensitivity in humans. There are limited data to directly compare animal and human sensitivity to OP compounds, although the mechanism is considered similar. There was no RBC cholinesterase inhibition in a single-dose study of humans at malathion doses as high as 15 mg/kg (Giles and Dickson 2000). This is higher than the 7.6 mg/kg estimate (95th percentile lower limit of the BMD_{10}) for malathion-induced RBC cholinesterase inhibition based on an acute dose to rats (U.S. EPA 2009b). Timchalk et al. (2002) developed physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) models for rats and humans for chlorpyrifos and found similar differences in chlorpyrifos sensitivity between rats and humans for RBC AChE inhibition. Even with a 100-fold uncertainty factor, the estimated AChE inhibition levels are low. At a DAP urinary level of 2000 nmol/L and assuming a 100-fold uncertainty factor, the AChE inhibition is estimated to be 0.2% for malathion, 2.1% for chlorpyrifos, 2.4% for dimethoate, and 21% for omethoate. While the omethoate estimate is above 10%, it was derived conservatively by assuming that all DAPs come from omethoate consumption, in addition to the 100-fold safety factor.

The available dose–response models are for acute exposures. There is no biological basis to determine whether the possible effects found in the epidemiologic studies are caused by acute (during a small window of pregnancy) or chronic (over the course of pregnancy) exposures. The dose–response analysis was done with acute exposures, because the DAP urinary measurements correspond to short-term exposures. Bradman et al. (2013) showed that there is a large variability in DAP measurements for individuals over one week. Thus, urinary DAP levels may not be appropriate for chronic dose–response assessment, unless steady state has been reached between dose rate and biotransformation/elimination, resulting in a plateau steady-state level of metabolite(s).

While most agree that OP toxicity is mediated through AChE inhibition, some have argued that toxicity from OP insecticides occurs at doses lower than those required to cause AChE inhibition (e.g., Slotkin and Seidler 2007). However, for many studies that have reached this conclusion, subsequent observations indicate that the AChE activity measurements from the inhibition tests were conducted long after the initial exposure. This allowed time for the AChE activity to recover, missing the point of maximum inhibition, and resulting in an underestimate of AChE inhibition (Eaton et al. 2008). Many of these studies were done with chlorpyrifos. It was also noted that the doses used in several of these studies ranged from 1 to 5 mg/kg chlorpyrifos administered subcutaneously to rat pups, or prenatally (Eaton et al. 2008). For 20 mL/kg/day of urine volume (Gonzales and Bauer 1999), assuming that 38% of DAPs are from exposure to chlorpyrifos (Zhang et al. 2008), the estimated DAP levels associated with 1–5 mg/kg of chlorpyrifos dose are approximately 375 000–1 900 000 nmol/L, levels that are well above those measured in the epidemiologic studies discussed in this section.

Some recent studies have also pointed to OP-mediated enzyme inhibition in the endocannabinoid system, which is important in nervous system development, and suggested that these effects occurred at doses that do not cause AChE inhibition (e.g., Carr et al. 2013). However, at this time, the meaning of these effects is unclear.

Overall, there are no toxicological data to suggest that deleterious effects could occur as a result of the low-level OP insecticide exposures experienced by subjects in the epidemiologic studies.

Confounding and bias

OP exposure in non-occupationally exposed populations is likely driven by diet and residential pesticide use (Krieger et al. 2012). Both diet/nutritional status and residential pesticide use may, in turn, be associated with other factors that affect health, thereby potentially resulting in confounding bias. In addition, selection bias can occur if study completion rates (in cohort studies) or participation rates (especially in case–control and cross-sectional studies) vary according to OP exposure, and health outcome.

For example, maternal body mass index (BMI), smoking, and nutrition can influence urinary DAP levels (see Figure 1 developed from CDC 2014 data; other data from CDC 2014 show that smokers have lower urinary DAP levels), as well as birth outcomes (Marshall and Spong 2012, Mason et al. 2012, Andres and Day 2000). These factors, along with childhood nutrition and BMI, which is inversely associated with urinary DAP levels (see Figure 2 developed from CDC 2014 data), can also influence neurodevelopmental outcomes in children

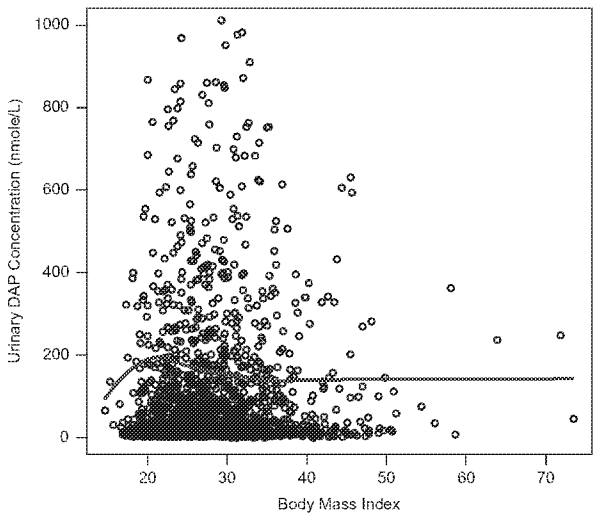


Figure 1. Urinary DAP (nmol/L) versus BMI for adults (>18 years of age) in the 2007–2008 NHANES dataset. Note: Graph truncated at 1000 nmol/L DAP concentration, which is about the 97th percentile. Red line produced with a LOESS smoothing function in the R programming language (R Core Team, 2014).

(Bliddal et al. 2014, Sandjaja et al. 2013, Neggers et al. 2003, Burkhalter and Hillman 2011, Anjos et al. 2013).

Another issue is that PON1, an enzyme that detoxifies some OP insecticides and that could therefore play an important role in mediating their toxic effects, may influence health outcomes independently of its effects on bioavailable OP levels—for example, through an antioxidant mechanism (Macharia et al. 2014). PON1 activity has a myriad of endogenous and environmental influences, including diet and lifestyle, as well as genetic determinants (Aviram and Vaya 2013, Schrader and Rimbach 2011). Thus, PON1 activity level could also confound apparent associations between DAP levels and health outcomes through a DAP-independent pathway.

In summary, numerous environmental and endogenous factors can affect birth outcomes and neurodevelopment, and many of these factors—including PON1 activity levels—may also influence OP internal dose and DAP levels, thereby leading to confounding. Selection bias may also occur if these factors influence study participation or completion rates. The full scope of determinants of OP and DAP exposure levels and of birth and neurodevelopmental outcomes is not known, and potentially vast. Even if statistical models adjust for several behavioral factors, residual confounding may occur due to omission of important variables or imprecise classification of those that are included.

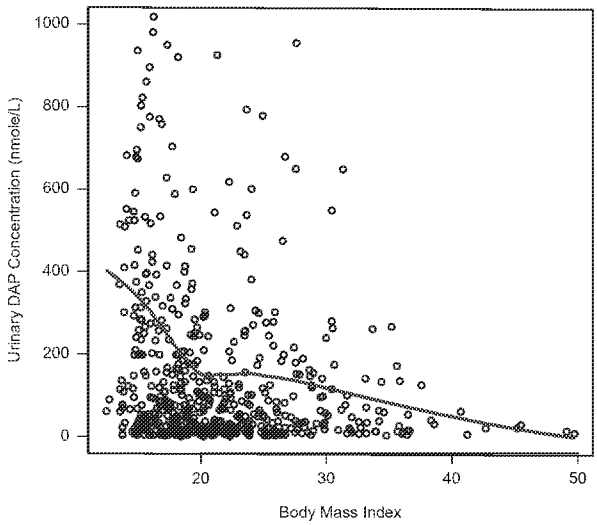


Figure 2. Urinary DAP (nmol/L) versus BMI for children (<19 years of age) in the 2007–2008 NHANES dataset. Note: Graph truncated at 1000 nmol/L DAP concentration, which is about the 97th percentile. Red line produced with a LOESS smoothing function in the R programming language (R Core Team, 2014).

Review of epidemiologic studies

Birth outcomes

Eleven studies in seven birth cohorts have investigated associations between OP metabolites and birth outcomes (Barr et al. 2010, Berkowitz et al. 2004, Eskenazi et al. 2004, Harley et al. 2011, Perera et al. 2003, Rauch et al. 2012, Wang et al. 2012, Whyatt et al. 2005, Whyatt et al. 2004, Wickerham et al. 2012, Wolff et al. 2007) (Table 1). All studies evaluated OP or OP metabolite levels in maternal prenatal or perinatal biospecimens and/or umbilical cord blood, in relation to standard measures of size and gestational age at birth ascertained from medical records, a computerized hospital perinatal database, and/or hospital delivery logs. Table 2 summarizes the analyses in the studies evaluating birth outcomes.

Columbia Center for Children's Environment and Health

The first study, based at the Columbia Center for Children's Environment and Health (CCCEH), followed healthy, nonsmoking, pregnant Dominican and African American women who had lived for at least one year in northern Manhattan or the South Bronx, New York, from ≤20 weeks of gestation through delivery (Table 1) (Perera et al. 2003, Whyatt et al. 2005, Whyatt et al. 2004). Study enrollment took place between 1998 and 2006. Chlorpyrifos, diazinon, and other pesticides were measured in maternal plasma samples collected within two days postpartum and in umbilical cord blood collected at delivery. Over the study period, average OP insecticide metabolite concentrations progressively declined. The mean concentration of chlorpyrifos was 7.1 pg/g in maternal plasma and 7.6 pg/g in cord plasma in an earlier study (Perera et al. 2003), but fell to 3.9 pg/g (standard deviation [SD] = 4.8) in maternal plasma and 3.7 pg/g (SD = 5.7) in cord plasma with extended enrollment (Whyatt et al. 2005). In the latter study, the mean concentration of diazinon was 1.3 pg/g (SD = 1.8) in maternal plasma and 1.2 pg/g (SD = 1.4) in cord plasma. OP insecticide levels were also measured in personal ambient air samples collected by mothers, who were asked to wear a backpack air sampling pump during the day and to place the monitor near the bed at night for two consecutive days during the third trimester of pregnancy. Mean air concentrations were 14.3 ng/m^3 (SD = 30.7) for chlorpyrifos and 99.5 ng/m^3 (SD = 449.8) for diazinon (Whyatt et al. 2005).

In multivariate adjusted linear regression models based on 263 mother–newborn pairs and with natural logarithm (ln)-transformed outcomes, maternal perinatal plasma chlorpyrifos levels (pg/g) were significantly inversely associated with birth weight (beta = −0.04 ln-g, $P = 0.01$) and birth length (beta = −0.03 ln-cm, $P = 0.04$), but not head circumference

Table 1. Design of epidemiologic studies of organophosphorus insecticide biomarkers.

Reference(s)	Study name	Location	Study design	Study subjects	Study dates	Exposure assessment	Exposure concentrations*	Outcome assessment
Perera et al. (2003), Whyatt et al. (2004, 2005), Rauh et al. (2006, 2011, 2012), Lovasi et al. (2011), Horton et al. (2012)	Columbia Center for Children's Environmental Health	New York City, New York, United States	Prospective birth cohort	Pregnant Dominican and African-American women aged 18–35 years, residing for ≥ 1 year before pregnancy in Washington Heights, Central Harlem, or South Bronx, New York, registered at one of two obstetrics and gynecology clinics by the 20th week of pregnancy, and without diabetes, hypertension, known HIV, or current use of tobacco or illicit drugs; 725 mother-child pairs enrolled, with 70% participation as of 2002; 83% retention rate at 3-year follow-up, 82% retention rate at 7-year follow-up. Rauh et al. (2012) further restricted to children with no/very low prenatal environmental tobacco smoke exposure and low prenatal airborne polycylic aromatic hydrocarbon exposure	1998–2006 up to age 7–11 years	Chlorpyrifos and diazinon (and other pesticides) measured in maternal plasma collected within 2 days postpartum and umbilical cord plasma collected at delivery; regression-derived maternal values were used in analyses when cord levels were unavailable Chlorpyrifos, diazinon, malathion, and methyl parathion (and other pesticides) measured with personal air monitor worn during daytime hours for 2 consecutive days and placed near bed at night during third trimester of pregnancy Chlorpyrifos and diazinon levels combined by converting diazinon levels to chlorpyrifos levels based on ratio of relative potency factors (6–1 for chlorpyrifos to diazinon) calculated by the U.S. Environmental Protection Agency (2002) Questionnaire administered at home during third trimester and annually thereafter	Maternal perinatal plasma (pg/g): Chlorpyrifos (Perera et al. 2003): mean = 7.1, 98% detectable Chlorpyrifos (Whyatt et al. 2005): mean ± SD = 3.9 ± 4.8 Diazinon (Whyatt et al. 2005): mean ± SD = 1.3 ± 1.8 Umbilical cord plasma (pg/g): Chlorpyrifos (Perera et al. 2003): mean = 7.6, 94% detectable Chlorpyrifos (Whyatt et al. 2005): mean ± SD = 3.7 ± 5.7 Diazinon (Whyatt et al. 2005): mean ± SD = 1.2 ± 1.4 Maternal prenatal personal air (ng/m^3) (Whyatt et al. 2005): Chlorpyrifos: mean ± SD = 14.3 ± 30.7 Diazinon: mean ± SD = 99.5 ± 449.8 Spearman correlation for maternal and cord plasma levels of chlorpyrifos = 0.6, $P < 0.001$ (Perera et al. 2003) and 0.79, $P \leq 0.001$ (Whyatt et al. 2005); diazinon = 0.69, $P \leq 0.001$ (Whyatt et al. 2005) Spearman correlation for maternal plasma and maternal air levels of chlorpyrifos = 0.21, $P \leq 0.001$; diazinon = 0.004, P-value NR (Whyatt et al. 2005)	Birth outcomes information and pregnancy and delivery characteristics obtained from mothers' and infants' medical records following delivery Bayley Scales of Infant Development, 2nd Edition (Mental Development Index and Psychomotor Development Index), administered at 12, 24, and 36 months Child Behavior Checklist for ages 1.5–5 years, including syndrome scale scores, internalizing and externalizing scores, and *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*-oriented scales, completed at 36 months Child Behavior Checklist for ages 6–18 years and Wechsler Intelligence Scale for Children, 4th Edition (Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index, and Processing Speed Index, combined for Full-Scale Intelligence Quotient), completed at 7 years Brain morphology assessed using high-resolution, T1-weighted magnetic resonance imaging at 5.9–11.2 years

Berkowitz et al. (2004), Engel et al. (2007), Wolff et al. (2007), Engel et al. 2011	Mount Sinai Children's Environmental Cohort Study	New York City, New York, United States	Prospective birth cohort	Consecutive primiparous pregnant women entering prenatal care with a singleton pregnancy at ≤ 26 weeks of gestation, without serious chronic disease or serious pregnancy complication, not consuming > 2 alcohol beverages per day or using illegal drugs, in a multi-ethnic, urban population; excluding infants with congenital malformation or severe prematurity (< 1,500 g or < 32 weeks of gestation); 479 (33%) participants of 1,450 eligible women; 404 included in analysis after excluding 75 (16%) of 479 due to medical complications, prematurity, congenital defect, lack of prenatal specimens, change of hospital or residence, or refusal (lower follow-up for younger and less-educated mothers)	1998–2001 up to age 6–9 years	TCPy, MDA, and six DAP metabolites (DMPs: dimethylphosphate, dimethylthiophosphate, and dimethyldithiophosphate; DEPs: diethylphosphate, diethylthiophosphate, and diethyldithiophosphate) measured in maternal urine collected during third trimester *PON1*$_{192}$, *PON1*$_{55}$, *PON1*$_{-909}$, *PON1*$_{-162}$, and *PON1*$_{-108}$ genotypes, PON1 activity (measured against phenylacetate) and butyrylcholinesterase activity (measured against butyrylthiocholine) assessed in third-trimester maternal blood and umbilical cord blood Prenatal questionnaire administered during third trimester	Median (IQR for TCPy; range for others) in maternal prenatal urine (Berkowitz et al. 2004, Wolff et al. 2007): TCPy: 7.6 (1.6–32.6) µg/L, 11.5 (1.8–35.4) µg/g creatinine MDA: limit of detection (< 0.3 µg/L) (< 0.3–15.8); 20.5% detectable DAPs: 75.9 (0–4,987) nmol/L, 88.6 (0–2,106) nmol/g creatinine DMPs: 42.2 (0–4,903) nmol/L, 55.4 (0–2,071) nmol/g creatinine DEPs: 18.8 (0–429) nmol/L, 22.1 (0–1,002) nmol/g creatinine	Birth outcomes information and delivery characteristics obtained from hospital computerized perinatal database Brazelton Neonatal Behavioral Assessment Scale administered before hospital discharge (age ≤ 5 days) and scored according to seven clusters developed by Lester et al. Bayley Scales of Infant Development, 2nd Edition (Mental Development and Psychomotor Development Indices), administered at ~12 and 24 months Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition (if age < 7 years), or Wechsler Intelligence Scale for Children, 4th Edition (if age 7–9 years) administered between ages 6 and 9 years

(*Continued*)

Table 1. (Continued)

Reference(s)	Study name	Location	Study design	Study subjects	Study dates	Exposure assessment	Exposure concentrations*	Outcome assessment
Eskenazi et al. (2004, 2007, 2010), Young et al. (2005), Marks et al. (2010), Bouchard et al. (2011), Harley et al. (2011), Quiros-Alcala et al. (2011)	Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS)	Salinas Valley, California, United States	Prospective birth cohort	Pregnant women aged ≥ 18 years, entering prenatal care at < 20 weeks of gestation, English- or Spanish-speaking, eligible for Medi-Cal, planning to deliver at the county hospital, in a primarily Latino, low-income, farmworker population; 601 (53.2%) participants of 1,130 eligible; ~530 [variably reported] followed through delivery of a live-born infant	1999–2000 up to age 7 years	Six DAP metabolites measured in maternal and child spot urines collected at interviews: dimethylphosphate, dimethylthiophosphate, and dimethyldithiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophosphate (combined as DEPs) Seven pesticide-specific metabolites measured in maternal spot urines collected at interviews: MDA, PNP, TCPy; also 2-diethylamino-4-hydroxy-6-methylpyrimidine, 2-isopropyl-4-methyl-6-hydroxypyrimidine, 3-chloro-4-methyl-7-hydroxycoumarin, and 5-chloro-1-isopropyl-3-hydroxytriazole (detectable in < 11%) Cholinesterase and butyrylcholinesterase measured in maternal blood/plasma taken at second interview during pregnancy and before delivery and in umbilical cord blood/plasma Maternal, cord, and child blood specimens genotyped for $PON1_{192}$ and $PON1_{-108}$; maternal post-delivery, umbilical cord, and 24-month child blood samples tested for PON1 enzyme quantity (arylesterase activity against phenylacetate) and enzyme activity (paraoxonase activity against paraoxon) Interviews at ~13–14 weeks of gestation, ~26–27 weeks of gestation, ~1 week after delivery, and when children were ~6 and 12 months and 2, 3.5, 5, and 7 years old	Median (range) in maternal urine (Eskenazi et al. 2004; Young et al. 2005): DAPs (nmol/L): 136 (10–6,854) prenatal, 222 (7–21,867) post-delivery DMPs (nmol/L): 101 (5–6,587) prenatal, 160 (5–21,857) post-delivery DEPs (nmol/L): 22 (2–680) prenatal, 27 (2–666) post-delivery MDA (μg/L): 0.2 (0.2–28.9) prenatal TCPy (μg/L): 3.3 (0.2–56.1) prenatal PNP (μg/L): 0.5 (0.1–34.7) prenatal Geometric mean (95% CI) in child urine (nmol/L) (Eskenazi et al. 2007; Marks et al. 2010): DAPs: 45.5 (39.6–52.3) at 6 months, 59.5 (51.7–68.5) at 12 months, 70.9 (61.4–81.9) at 24 months, 77.5 (65.4–91.9) at 3.5 years, 92.6 (87.6–109.0) at 5 years DMPs: 23.8 (20.4–27.8) at 6 months, 32.9 (27.8 38.9) at 12 months, 48.6 (41.8–56.6) at 24 months, 62.5 (52.2–74.7) at 3.5 years, 72.4 (61.0–86.0) at 5 years DEPs: 10.6 (8.9–11.9) at 6 months, 15.2 (13.5–17.2) at 12 months, 10.5 (8.8–12.6) at 24 months, 7.0 (5.8–8.3) at 3.5 years, 7.2 (6.0–8.7) at 5 years	Birth outcomes information obtained from hospital delivery logs and medical records Brazelton Neonatal Behavioral Assessment Scale administered at or before 62 days, with seven clusters developed by Lester et al. Bayley Scales of Infant Development, 2nd Edition (Mental Development and Psychomotor Development Indices), administered at 6, 12, and 24 months Autonomic nervous system reactivity protocol administered using social, physical, emotional, and cognitive (at 3.5 and 5 years) challenges, with measurement of heart rate, respiratory sinus arrhythmia, and pre-ejection period at 6 months and 1, 3.5, and 5 years Child Behavior Checklist for ages 1.5–5 years (attention problems syndrome, ADHD, and pervasive developmental disorder scales) completed by mothers at 2, 3.5, and 5 years NEPSY®-II visual attention subtest administered at 3.5 years Conners' Kiddie Continuous Performance Test (for reaction time, accuracy, and impulse control) administered and Hillside Behavior Rating Scale (for motor activity and distractibility) completed by psychometricians at 5 years Wechsler Intelligence Scale for Children, 4th edition, administered at 7 years

Lizardi et al. 2008	Children Pesticide Survey	Yuma County, Arizona, United States	Cross-sectional	Schoolchildren (mean age = 7 years) from an agricultural community near the U.S.-Mexico border, previously participating in a pesticide screening study and selected for further study based on the absence ($N = 23$) or presence ($N = 28$) of urinary organophosphate pesticide metabolites in the original urine specimen	2002	Six DAP metabolites measured in first-void urine sample collected from each child on the day of the cognitive assessment: dimethylphosphate, dimethylthiophosphate, dimethyldithiophosphate, diethylphosphate, diethylthiophosphate, and diethyldithiophosphate; lower limit of detection = 25 µg/L Structured interview at home with parents	Mean ± SD dimethylphosphate (µg/L) in child urine = 65.5 ± 78 (95% CI = 43, 88), 100% detectable Mean (95% CI) urinary DAPs (µg/L) in originally exposed group = 110 (83–139); mean in originally unexposed group = 49 (36–63), $P < 0.01$, after excluding one high outlier from each group (In original screening urine sample, 25 children had detectable and 23 had undetectable DAPs)	Cognitive performance assessed using Wechsler Intelligence Scale for Children–Third Edition Short Form, Children's Memory Scale, Wisconsin Card Sorting Test, and Trail Making Test A and B, completed by child at school or, if not possible, at home during a second visit Behavioral performance assessed using Child Behavior Checklist/4–18 (completed at home by parents) and Teacher Report Form (completed at school by teachers)
Barr et al. (2010)	–	New Jersey, United States	Prospective birth cohort	Convenience sample of 150 women with a singleton pregnancy and non-anomalous fetus scheduled for an elective cesarean birth at term (≥ 37 weeks of gestation) with hemoglobin level ≥ 8 mg/dL, excluded if evidence for labor or rupture of membranes at time of operative delivery or if using medications that could potentially interfere with metabolism or environmental chemicals; 2 maternal blood and 2 umbilical cord blood samples excluded due to processing errors	2003–2004 to birth	Chlorpyrifos and other pesticides measured in maternal blood obtained prior to placement of intravenous and bladder catheters before cesarean section or in extra maternal blood specimens available from preoperative testing Also measured in umbilical cord blood obtained within 15 minutes of delivery Self-administered questionnaire distributed to pregnant women	Chlorpyrifos in maternal serum (ng/g): 98.6% detectable, mean ± SD = 0.09 ± 0.87, median (range, IQR) = 0.0007 (0.0007–10.09, 0.0007–0.0007) Chlorpyrifos in umbilical cord serum (ng/g): 62.8% detectable, mean ± SD = 0.55 ± 0.73, median (range, IQR) = 0.0007 (0.0007–1.84, 0.0007–1.32) Pearson's correlation for chlorpyrifos in maternal and cord serum = 0.12	Birth outcomes information and pregnancy characteristics obtained from medical records prior to hospital discharge

(*Continued*)

Table 1. (Continued)

Reference(s)	Study name	Location	Study design	Study subjects	Study dates	Exposure assessment	Exposure concentrations*	Outcome assessment
Bouchard et al. (2010)	National Health and Nutrition Examination Survey (NHANES) 2000–2004	United States	Cross-sectional	Population-based health survey data from non-institutionalized children aged 8–15 years selected using multi-stage probability sampling, with oversampling of certain subgroups, to be representative of the general U.S. population; ADHD assessed in 3,998 participants, urinary DAP metabolite data available for 1,481 (37%) based on 50% sampling rate for ages 6–11 years and 33% for ages 12–15 years in 2000–2002, and 33% sampling rate in 2003–2004; further excluded children who received NICU or premature nursery care and those with birth weight < 2,500 g, urinary creatinine < 20 mg/dL, outlier urinary DAP concentrations, or missing covariate data	2000–2004	Six DAP metabolites measured in spot urine samples collected during physical examinations at mobile study centers: dimethylphosphate, dimethylthiophosphate, dimethyldithiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophosphate (combined as DEPs) Interview conducted in respondent's home	Geometric mean (range and IQR) in child urine (nmol/L): DAPs: 68.3 (6.0–10,195, 24.4–186.0) DMPs: 41.3 (4.5–10,068, 10.1–130.7) DEPs: 11.0 (0.8–5905, 2.1–35.0) Dimethylphosphate: 10.7 (2.8–1324, 2.8–39.0) Dimethylthiophosphate: 13.7 (0.9–9929, 1.9–58.8) Dimethyldithiophosphate: 1.7 (0.3–7006, 0.4–7.3) Diethylphosphate: 4.7 (0.4–5902, 0.9–28.1) Diethylthiophosphate: 2.0 (0.3–650, 0.4–7.6) Diethyldithiophosphate: 0.5 (0.2–36, 0.3–0.3)	ADHD and ADHD subtypes in previous year assessed based on Diagnostic Interview Schedule for Children IV based on slightly modified criteria from *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, based on a telephone interview with the mother or another caretaker 2–3 weeks after physical examination ADHD defined in study as meeting diagnostic criteria of ADHD or regularly taking ADHD medication during the previous year
Wang et al. (2011)	–	Shanghai, China	Prospective birth cohort	Pregnant women aged 18–45 years attending one of two major obstetric hospitals, with no gestational or pre-existing diabetes, hypertension, HIV/AIDS, or use of illegal drugs in the preceding year, with singleton infants free of severe neonatal illness; 187 (96.9%) participants of 193 eligible	2006–2007 to birth	Five DAP metabolites measured in maternal spot urines collected at the onset of labor: dimethylphosphate, dimethylthiophosphate, diethylphosphate, diethylthiophosphate, and diethyldithiophosphate Interview conducted during pregnancy	Geometric mean (range and IQR) in maternal prenatal urine (μg/L): Dimethylphosphate: 17.19 (undetectable–269.15; 7.02–53.70) Dimethylthiophosphate: 8.01 (undetectable–109.65; 3.53–20.06) Diethylphosphate: 6.03 (undetectable–109.65; 3.55–11.17) Diethylthiophosphate: 6.31 (undetectable–131.83; 3.36–11.98) Diethyldithiophosphate: NR because 5.34% detectable (undetectable–5.1; undetectable–undetectable) Geometric mean (range and IQR) in maternal prenatal urine (μg/g creatinine): Dimethylphosphate: 25.75 (0.81–588.84; 12.25–72.86) Dimethylthiophosphate: 11.99 (0.56–123.02; 5.45–28.40) Diethylphosphate: 9.03 (0.58–89.13; 5.13–16.54) Diethylthiophosphate: 9.45 (0.47–93.33; 4.53–18.30) Diethyldithiophosphate: NR because 5.34% detectable (0.31–9.33; 0.94–2.43)	Birth outcomes information and pregnancy and delivery characteristics obtained from mothers' and infants' medical records

Guodong et al. (2012)	–	Shanghai, China	Cross-sectional	Children aged 23–25 months attending routine physical check-ups at departments of child and adolescent health care at two community hospitals, with no intrauterine distress, pathological jaundice, intrauterine infection, intracranial infection, congenital disease, or current cold or fever, and able to complete the neurodevelopmental assessment; 301 (97.1%) participants of 310 eligible	2008	Five DAP metabolites measured in spot urine collected on the day of study assessment: dimethylphosphate, dimethylthiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophosphate (combined as DEPs) Interview conducted with mothers	Geometric mean (range and IQR) in child urine (μg/L): Dimethylphosphate: 2.52 (< 2.0 [limit of detection]–186.99; < 2.0–3.41) Dimethylthiophosphate: 1.56 (< 1.0–80.81; < 1.0–1.63) Diethylphosphate: 1.78 (< 1.0–32.19; < 1.0–2.89) Diethylthiophosphate: 3.18 (< 1.0–55.40; < 1.0–7.26) Diethyldithiophosphate: NR because 2.7% detectable (< 1.0–3.80; < 1.0–< 1.0) Geometric mean (range and IQR) in child urine (μg/g creatinine): Dimethylphosphate: 11.27 (1.53–729.27; 4.33–24.02) Dimethylthiophosphate: 6.99 (1.08–481.50; 3.09–13.12) Diethylphosphate: 7.96 (1.14–170.96; 3.84–16.36) Diethylthiophosphate: 14.19 (1.10–980.58; 5.30–37.15) Diethyldithiophosphate: 4.55 (1.08–73.14; 2.49–7.70)	Gesell Developmental Schedules for 0- to 3-year-old children administered to evaluate neurological and intellectual development using four main categories of functioning: motor behavior, adaptive behavior, language behavior, and personal and social behavior
Rauch et al. (2012), Yolton et al. (2013)	Health Outcomes and Measures of the Environment (HOME) Study	Cincinnati, Ohio, United States	Prospective birth cohort	Pregnant women aged ≥ 18 years attending one of seven prenatal clinics, living in a home built before 1978, ≤ 19 weeks of gestation, HIV-negative, living within five surrounding counties in a socioeconomically diverse area, and not receiving thyroid or seizure medications, or chemotherapy or radiation treatments; 468 (37.1%) participants of 1,263 eligible; 389 followed through delivery of a live-born singleton infant (9 followed through delivery of twins, 3 followed through stillbirth)	2003–2006 up to age ~5 weeks	Six DAP metabolites measured in maternal spot urines collected at ~16 and ~26 weeks of gestation (averaged for analysis) and within 24 hours of delivery: dimethylphosphate, dimethylthiophosphate, and dimethyldithiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophospate (combined as DEPs) Umbilical cord blood genotyped for $PON1_{192}$ and $PON1_{-108}$ Interviews during pregnancy (~20 weeks of gestation) and after delivery (~5 weeks)	Median (IQR) in maternal prenatal urine (nmol/L) (Rauch et al. 2012): DAPs: 81.3 (41.7–220.0) DMPs: 56.9 (26–185) DEPs: 17.7 (8–37) 10-fold increase in DAPs ≈ 15th percentile (29.5 nmol/L) to 85th percentile (318.0 nmol/L)	Birth weight abstracted from medical records; gestational age calculated from mother's self-reported date of last menstrual period or based on ultrasound ($N = 7$) or Ballard examination at delivery ($N = 3$) NICU Network Neurobehavioral Scale administered in home at ~5 weeks (mean = 34 days, range = 17–47); 13 dimensions: habituation (omitted due to small number completed), attention, arousal, self-regulation, need for special handling by examiner, quality of movement, excitability, lethargy, non-optimal reflexes, asymmetrical reflexes, hypertonicity, hypotonicity, and stress/abstinence

(*Continued*)

Table 1. (Continued)

Reference(s)	Study name	Location	Study design	Study subjects	Study dates	Exposure assessment	Exposure concentrations*	Outcome assessment
Wickerham et al. (2012)	–	Zhejiang, China	Prospective birth cohort	Consecutive pregnant women with a healthy, uncomplicated, singleton pregnancy recruited from a single hospital at 36 weeks of gestation, excluding those with chronic diseases, complicated pregnancies, or hereditary or metabolic diseases; 116 participants with infants born at > 37 weeks of gestation and umbilical cord blood pesticide levels (~99.6% participation rate); excluded 3 with missing data and 1 highly influential outlier	2009 to birth	Eight organophosphate pesticides (and other pesticides) measured in umbilical cord serum at delivery: chlorpyrifos, diazinon, fonofos, malathion, parathion-ethyl, parathion-methyl, profenofos, and terbufos	% detectable, median, and 90th percentile in umbilical cord serum at delivery (ng/mL): Chlorpyrifos: 23.3%, < 0.05 (limit of detection), 0.17 Diazinon: 14.7%, < 0.05 (limit of detection), 0.27 Fonofos: 16.4%, < 0.05 (limit of detection), 0.30 Malathion: 25.9%, < 0.50 (limit of detection), 3.13 Parathion-ethyl: 2.6%, < 0.05 (limit of detection), < 0.05 Parathion-methyl: 28.5%, < 0.05 (limit of detection), 1.43 Profenofos: 25.0%, < 0.50 (limit of detection), 0.68 Terbufos: 31.0%, < 0.05 (limit of detection), 0.27	Birth outcomes information and pregnancy characteristics obtained from patient charts
Oulhote and Bouchard (2013)	Canadian Health Measures Survey, cycle 1	Canada	Cross-sectional	Population-based health survey data from children selected using multi-stage probability sampling, with oversampling of certain subgroups, to be representative of the general Canadian population; 1,081 children aged 6–11 years among 5,600 participants aged 6–79 years; 1,030 (95%) with most urinary pesticide metabolite levels and behavioral assessment, 779 (72%) after exclusion of those with missing covariate data	2007–2009	Six DAP metabolites measured in spot urine samples collected during physical examinations at mobile examination centers within 2 weeks of survey questionnaire completion: dimethylphosphate, dimethylthiophosphate, dimethyldithiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophosphate (combined as DEPs) Household survey conducted in respondent's home	Median (IQR) in child urine (nmol/L): DAPs: 99.2 (34.3–273.3) DMPs: 62.0 (18.7–192.8) DEPs: 25.0 (10.5–51.3) Dimethylphosphate: 34.6 (10.8–91.9) Dimethylthiophosphate: 17.6 (< 4.2 [limit of detection]–75.4) Dimethyldithiophosphate: < 1.9 (limit of detection) (< 1.9–5.6) Diethylphosphate: 19.6 (8.5–42.0) Diethylthiophosphate: < 3.5 (limit of detection) (< 3.5–6.9) Diethyldithiophosphate: < 1.6 (limit of detection) (< 1.6–< 1.6)	Behavioral problems assessed using parent version of the Strengths and Difficulties Questionnaire, including scales for emotional symptoms, conduct problems, hyperactivity/inattention, peer problems, prosocial behavior, and total difficulties (sum of all dimension scales except prosocial behavior), categorized into high vs. low/borderline using author-recommended cutoff scores

Fortenberry et al. (2014)	Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT)	Mexico City, Mexico	Prospective birth cohort	Mother-child pairs from three sequentially-enrolled cohorts of pregnant women enrolled during pregnancy or at delivery from a general hospital or affiliated clinics in a low- to moderate-income setting, excluding women with plans to leave the area within five years, daily alcohol consumption, addiction to illegal drugs, continuous use of prescription drugs, diagnosis of multiple pregnancy, pre-eclampsia, renal or heart disease, gestational diabetes, high-risk pregnancy, or seizures requiring medical treatment, or history of infertility, diabetes, or psychosis; 187 (23%) of 827 participants re-invited from second and third cohorts with child psychometric assessment and third-trimester maternal urine, including 21 with urine in all three trimesters	1994–1997, 1997–2000, or 2001–2005–2007–2011 (ages 6–11 years)	TCPy measured in maternal third-trimester morning void urine specimens	Geometric mean (95% CI and IQR) TCPy in maternal prenatal urine (ng/mL) = 1.76 (1.55–2.02, 0.91–3.57) Intraclass correlation among 21 subjects with measured levels in all three trimesters of pregnancy = 0.41 without correction for specific gravity, 0.29 with correction	Conners' Parent Rating Scales–Revised (ADHD Index, Global Restlessness/Impulsivity Index, and Hyperactivity/Impulsivity, Inattention, and Combined ADHD scales based on guidelines from *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*) completed by parents Behavior Assessment System for Children–Parental Rating Scales completed by parents Conners' Continuous Performance Test completed by children
Zhang et al. (2014)	–	Shenyang, China	Prospective birth cohort	Healthy pregnant women recruited from a single hospital, living in Shenyang for > 3 years, without hypertension, diabetes, thyroid hypofunction, heart disease, or other chronic diseases before pregnancy, without serious pregnancy complications, and without family or medical history of mental retardation, phenylketonuria, or Pompe's syndrome for self or spouse; also excluding infants with disorders associated with adverse neurodevelopment; 249 (81.1%) participants of 307 eligible	2011–2012 to age 3 days	Five DAP metabolites measured in maternal prenatal urine (timing not specified): dimethylphosphate, dimethylthiophosphate (combined as DMPs), diethylphosphate, diethylthiophosphate, and diethyldithiophosphate (combined as DEPs) In-person interviews with mothers	Geometric mean (range and IQR) in maternal prenatal urine (μg/L): Dimethylphosphate: 18.03 (< 2 [limit of detection]–334.02, 7.83–39.43) Dimethylthiophosphate: 8.53 (< 1 [limit of detection]–137.95, 3.4–15.67) Diethylphosphate: 7.14 (< 1 [limit of detection]–167.06, 3.54–17.17) Diethylthiophosphate: 5.64 (< 1 [limit of detection]–133.00, 2.34–13.55) Diethyldithiophosphate: < 1 (limit of detection) (< 1–6.61, < 1–< 1) Geometric mean (range and IQR) in maternal prenatal urine (μg/g creatinine): Dimethylphosphate: 24.02 (0.19–453.04, 9.72–60.1) Dimethylthiophosphate: 11.29 (0.10–305.92, 4.07–29.49) Diethylphosphate: 9.49 (0.36–125.09, 4.29–20.45) Diethylthiophosphate: 7.58 (0.32–102.21, 2.93–19.95) Diethyldithiophosphate: 0.78 (0.03–15.41, 0.46–1.34)	Neonatal Behavioral Neurological Assessment performed at age 3 days, with five scales: behavior, passive tone, active tone, primary reflexes, and general assessment, combined as summary score

*Values shown are reported in the earliest available publication from each study cohort, except for the Columbia cohort, where values changed substantially over time and are shown from multiple publications.

ADHD attention deficity/hyperactivity disorder confidence interval, *DAP* dialkyl phosphate, *DEP* diethyl phosphate, *DMP* dimethyl phosphate, *IQR* interquartile range, *MDA* malathion dicarboxylic acid, *NICU* neonatal intensive care unit, *NR* not reported, *PNP* 4-nitrophenol, *PON1* paraoxonase 1, *SD* standard deviation, *TCPy* 3,5,6-trichloro-2-pyridinol.

(beta = −0.005 ln-cm, $P = 0.82$) (Table 2) (Perera et al. 2003). The inverse association with birth weight was statistically significant among African Americans but not Dominicans, whereas the reverse race/ethnicity pattern was observed for birth length. In subsequent analyses based on 314 mother–newborn pairs, cord plasma chlorpyrifos levels (ln-pg/g) were also significantly inversely associated with birth weight (beta = −42.6 g, 95% confidence interval [CI] = −81.8, −3.8) and birth length (beta = −0.24 cm, 95% CI = −0.47, −0.01), but not head circumference (beta = −0.01 cm, 95% CI = −0.13, 0.11) (Whyatt et al. 2004). Slightly stronger inverse associations were observed with cord plasma chlorpyrifos and diazinon levels combined, but diazinon itself was not significantly associated with any of the three outcomes. Maternal prenatal personal air levels of chlorpyrifos, diazinon, and both OPs combined also were not significantly associated with any of the three birth outcomes. The inverse associations between cord plasma chlorpyrifos and birth weight and length were restricted to newborns born before January1, 2001, when the U.S. EPA instituted regulations to phase out residential use of these insecticides; exposure levels were substantially lower and no associations were detected in newborns born in 2001 or later. Similar findings were reported in a slightly larger group of mother–newborn pairs with cord plasma measures of chlorpyrifos and diazinon (Whyatt et al. 2005). Specifically, birth weight was 67.3 g lower (95% CI = −116.6, −17.8), and birth length was −0.43 cm shorter (95% CI = −0.73, −0.14) for each one-unit (ln-pg/g) increase in cord plasma chlorpyrifos among 237 newborns born before January 1, 2001, but no such association was detected among 77 newborns born after that date (beta for birth weight = 30.7 g, 95% CI = −108.6, 169.9; beta for birth length = 0.07 cm, 95% CI = −0.65, 0.79).

Substantial strengths of the CCCEH study include the use of objective, individually measured metabolites to characterize exposure to OP insecticides (a strength of all studies discussed in this review), the availability of information on numerous potential confounders, and the prospective design, with maternal interviews and personal air monitoring conducted during the third trimester of pregnancy, prior to the health outcomes of interest.

Some methodological limitations of the CCCEH study should be noted. First, a single maternal blood sample was collected from each subject at or shortly after delivery. Normal fetal growth is approximately linear between 18 and 37 weeks of gestation, after which it plateaus; thus, maternal plasma OP levels at delivery may not reflect levels in past weeks or months, and may be etiologically irrelevant to fetal growth. Although maternal air samples were obtained in the third trimester of pregnancy, it is unknown whether a single sample collected over two days is representative of exposure at other time points. Second, the number of participants was modest, especially after stratification by race/ethnicity or birth date, resulting in several statistically unstable estimates of association (i.e., wide confidence bounds). Third, given the many potential influences on chlorpyrifos and diazinon levels in peripheral blood and air, as well as on birth outcomes, uncontrolled confounding by diet and other factors may partially explain some of the observed results. However, without detailed knowledge of established predictors of chlorpyrifos and diazinon exposure and of birth outcomes in this study population, the direction of potential confounding is difficult to predict, and the magnitude is probably limited by the adjustment for several major influences on birth outcomes. Fourth, because numerous hypotheses were tested, at least some statistically significant associations are expected due to chance. Neither this study nor any other study of birth outcomes described in this review made statistical corrections for multiple comparisons. Although such corrections are not standard in traditional epidemiology, authors who do not correct for multiple comparisons should report the number and nature of all associations tested, how certain associations were selected for reporting, and the probable effect of such selection on the results (Rothman et al. 2012). As evidence in other areas of research, particularly genetic epidemiology, numerous exploratory analyses almost inevitably lead to false-positive results and recent methodological literature includes several practical ways of dealing with this problem (Wacholder et al. 2004, Strömberg et al. 2008, Weitkunat et al. 2010, Wakefield 2007).

Finally, the completeness of follow-up from enrollment through delivery was not reported, but if follow-up varied by uncontrolled factors, such as diet, that might be associated with maternal OP exposure and birth outcomes, then an unpredictable degree of selection bias could have occurred. Cohort participation rates also were not reported (but were stated as 70% in an earlier publication [Whyatt et al. 2002]), and could have been a source of a moderate degree of selection bias if participation were related to OP exposure and birth outcomes. (Participation bias is usually considered not to be a major concern in prospective cohort studies, because outcomes occur after cohort entry, but with relatively short-term follow-up, it is conceivable that participation could be associated with risk of adverse birth outcomes.)

In summary, results in the CCCEH cohort suggest an inverse association of maternal perinatal plasma levels of chlorpyrifos, but not diazinon, with birth weight and birth length, but not with head circumference, in an urban, low-income, minority population. The observation of associations only among newborns born before 2001, when exposure levels were higher, suggests a possible exposure threshold below which chlorpyrifos is not associated with birth outcomes. The detection of certain associations only in African Americans but not Dominicans, or vice versa, indicates that the observed associations may be attributable to excessive stratification. The lack of associations with maternal prenatal personal air levels of chlorpyrifos and diazinon raises the question of whether route of exposure is an effect modifier of associations between OP exposure and birth outcomes. Furthermore, the different results for chlorpyrifos and diazinon suggest that OP insecticides should be analyzed separately, not combined, with respect to birth outcomes, although this approach raises the problem of multiple comparisons. Overall, the inconsistent results by outcome, racial/ethnic group, and exposure metric render the findings difficult to interpret, and do not provide compelling evidence to support an adverse effect of chlorpyrifos or diazinon on fetal growth.

Mount Sinai Children's Environmental Cohort Study

Another birth cohort based in New York City, the Mount Sinai Children's Environmental Cohort Study (CECS), enrolled

Table 2. Results of epidemiologic studies of organophosphorus insecticide biomarkers and birth outcomes.

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Perera et al. (2003)	Birth weight (g), natural log scale	Maternal perinatal plasma chlorpyrifos (pg/g)	263 total 116 African American 146 Dominican	Beta = −0.04, $P = 0.01$ Beta = −0.05, $P = 0.04$ Beta = −0.02, $P = 0.26$	Maternal body mass index, parity, cotinine, infant sex, gestational age, and maternal prenatal airborne polycyclic aromatic hydrocarbon levels	No significant interactions were observed between chlorpyrifos and polycyclic aromatic hydrocarbons, although numbers were limited (results NR)
Perera et al. (2003)	Birth length (cm), natural log scale	"	263 total 116 African American 146 Dominican	Beta = −0.02, $P = 0.04$ Beta = −0.01, $P = 0.15$ Beta = −0.02, $P = 0.002$	"	–
Perera et al. (2003)	Head circumference (cm), natural log scale	"	263 total 116 African American 146 Dominican	Beta = −0.005, $P = 0.28$ Beta = −0.003, $P = 0.70$ Beta = −0.005, $P = 0.31$	"	–
Whyatt et al. (2004)	Birth weight (g)	Cord plasma chlorpyrifos (pg/g, natural log scale)	314 total 237 born before 1 January 2001 77 born before 1 January 2001	Beta = −42.6 (−81.8, −3.8) Birth before 1 January 2001 beta = −67.3 (−116.6, −17.8) Birth after 1 January 2001 beta = 30.7 (−108.6, 169.9) Group 2 vs. 1 beta = 39.2 (−107.3, 185.7) Group 3 vs. 1 beta = −50.9 (−188.2, 86.3) Group 4 vs. 1 beta = −150.1 (−287.7, −12.5)	Gestational age, maternal pre-pregnancy weight, maternal weight gain during pregnancy, newborn sex, parity, race/ethnicity, environmental tobacco smoke in home, and season of delivery No change after additional adjustment for cord plasma 2-isopropoxyphenol levels (results NR)	Except for plasma diazinon levels, insecticide levels decreased substantially for infants born after 1 January 2001 (after phase-out of residential use by U.S. Environmental Protection Agency regulatory action), despite no significant change in self-reported pesticide use habits Group 1: < limit of detection (31% of cord plasma chlorpyrifos levels, 48% of cord plasma diazinon levels); groups 2, 3, and 4: tertiles of detectable levels
Whyatt et al. (2004)	"	Cord plasma chlorpyrifos + diazinon (pg/g, natural log scale)	"	Beta = −49.1 (−91.3, −6.9) Birth before 1 January 2001 beta = −72.5 (−125.0, −20.0) Birth after 1 January 2001 beta = 0.6 (−144.7, 145.9) Group 2 vs. 1 beta = −78.5 (−225.5, 68.5) Group 3 vs. 1 beta = −33.1 (−173.7, 107.4) Group 4 vs. 1 beta = −186.3 (−327.2, −45.4)	"	Group 1: < limit of detection (31% of cord plasma chlorpyrifos levels, 48% of cord plasma diazinon levels); groups 2, 3, and 4: tertiles of detectable levels
Whyatt et al. (2004)	"	Cord plasma diazinon (pg/g, natural log scale)	"	Beta = −44.2 (−119.5, 31.0)	"	–
Whyatt et al. (2004)	"	Maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (ng/m^3, natural log scale)	"	Chlorpyrifos beta = −17.7 (−64.2, 28.9) Diazinon beta = 13.8 (−23.2, 50.8) Chlorpyrifos + diazinon beta = −5.1 (−50.7, 40.4)	"	Associations with maternal personal air samples remained non-significant after stratification by birth before or after 1 January 2001 (results NR)

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Whyatt et al. (2004)	Birth length (cm)	Cord plasma chlorpyrifos (pg/g, natural log scale)	309 total 237 born before 1 January 2001 77 born after 1 January 2001	Beta = −0.24 (−0.47, −0.01) Birth before 1 January 2001 beta = −0.43 (−0.73, −0.14) Birth after 1 January 2001 beta = 0.07 (−0.65, 0.79) Group 2 vs. 1 beta = 0.17 (−0.70, 1.0) Group 3 vs. 1 beta = −0.21 (−1.0, 0.61) Group 4 vs. 1 beta = −0.75 (−1.6, 0.06)	"	Group 1: < limit of detection (31% of cord plasma chlorpyrifos levels, 48% of cord plasma diazinon levels); groups 2, 3, and 4: tertiles of detectable levels
Whyatt et al. (2004)	"	Cord plasma chlorpyrifos + diazinon (pg/g, natural log scale)	"	Beta = −0.27 (−0.52, −0.02) Birth before 1 January 2001 beta = −0.46 (−0.77, −0.14) Birth after 1 January 2001 beta = −0.07 (−0.82, 0.67) Group 2 vs. 1 beta = −0.06 (−0.93, 0.81) Group 3 vs. 1 beta = −0.005 (−0.84, 0.82) Group 4 vs. 1 beta = −0.80 (−1.6, 0.02)	"	Group 1: < limit of detection (31% of cord plasma chlorpyrifos levels, 48% of cord plasma diazinon levels); groups 2, 3, and 4: tertiles of detectable levels
Whyatt et al. (2004)	"	Cord plasma diazinon (pg/g, natural log scale)	"	Beta = −0.32 (−0.75, 0.11)	"	–
Whyatt et al. (2004)	"	Maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (ng/m^3, natural log scale)	"	Chlorpyrifos beta = −0.02 (−0.28, 0.25) Diazinon beta = 0.07 (−0.14, 0.28) Chlorpyrifos + diazinon beta = −0.01 (−0.27, 0.25)	"	Associations with maternal personal air samples remained non-significant after stratification by birth before or after 1 January 2001 (results NR)
Whyatt et al. (2004)	Head circumference (cm)	Cord plasma chlorpyrifos, diazinon, or chlorpyrifos + diazinon (pg/g, natural log scale)	298 total	Chlorpyrifos beta = −0.01 (−0.13, 0.11) Diazinon beta = −0.07 (−0.30, 0.16) Chlorpyrifos + diazinon beta = −0.02 (−0.15, 0.11)	Gestational age, maternal pre-pregnancy weight, maternal weight gain during pregnancy, newborn sex, parity, race/ethnicity, environmental tobacco smoke in home, season of delivery, and cesarean section delivery No change after additional adjustment for cord plasma 2-isopropoxyphenol levels (results NR)	–

Whyatt et al. (2004)	"	Maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (ng/m^3, natural log scale)		Chlorpyrifos beta = − 0.04 (− 0.18, 0.10) Diazinon beta = − 0.03 (− 0.14, 0.09) Chlorpyrifos + diazinon beta = − 0.03 (− 0.17, 0.11)	"	Associations with maternal personal air samples remained non-significant after stratification by birth before or after 1 January 2001 (results NR)
Whyatt et al. (2005)	Birth weight (g)	Cord plasma chlorpyrifos (pg/g, natural log scale)	237 born before 1 January 2001 77 born after 1 January 2001	Beta = − 67.3 (− 116.6, − 17.8) Beta = 30.7 (− 108.6, 169.9)	Gestational age, maternal pre-pregnancy weight, maternal weight gain during pregnancy, newborn gender, parity, ethnicity, environmental tobacco smoke in home, and season of delivery No change after additional adjustment for cord plasma 2-isopropoxyphenol levels (results NR)	34% of newborns born before 1 January 2001 and 1.5% of those born after had cord plasma levels of chlorpyrifos + diazinon in the top tertile of detectable levels ($P < 0.001$)
Whyatt et al. (2005)	"	Cord plasma chlorpyrifos + diazinon (pg/g, natural log scale)	237 born before 1 January 2001 77 born after 1 January 2001	Beta = − 72.5 (− 125.0, − 20.0) Beta = 0.6 (− 144.7, 145.9) Group 4 vs. 1 beta = − 215.1 (− 384.7, − 45.5)	"	Group 1: < limit of detection; groups 2, 3, and 4: tertiles of detectable levels
Whyatt et al. (2005)	"	Maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (ng/m^3, natural log scale)	237 born before 1 January 2001 77 born after 1 January 2001	"No association" (results NR)	"	–
Whyatt et al. (2005)	Birth length (cm)	Cord plasma chlorpyrifos (pg/g, natural log scale)	237 born before 1 January 2001 77 born after 1 January 2001	Beta = − 0.43 (− 0.73, − 0.14) Beta = 0.07 (− 0.65, 0.79)	"	–
Whyatt et al. (2005)	"	Cord plasma chlorpyrifos + diazinon (pg/g, natural log scale)	237 born before 1 January 2001 77 born after 1 January 2001	Beta = − 0.46 (− 0.77, − 0.14) Beta = − 0.07 (− 0.82, 0.67)	"	–
Whyatt et al. (2005)	"	Maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (ng/m^3, natural log scale)	"	"No association" (results NR)	"	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Whyatt et al. (2005)	Head circumference (cm)	Cord plasma or maternal prenatal personal air chlorpyrifos, diazinon, or chlorpyrifos + diazinon (natural log scale)	"	"No association" (results NR)	Gestational age, maternal pre-pregnancy weight, maternal weight gain during pregnancy, newborn gender, parity, ethnicity, environmental tobacco smoke in home, season of delivery, and delivery by cesarean section	–
Berkowitz et al. (2004)	Birth weight (g)	Maternal prenatal urinary TCPy (μg/L)	216 < 11.0 μg/L (limit of detection) 171 > 11.0 μg/L	Mean ± SD = 3,284 ± 441 Mean ± SD = 3,296 ± 434 $P > 0.05$	Race/ethnicity, infant sex, and gestational age No difference after additional adjustment for active and passive cigarette smoking, pre-pregnancy body mass index, maternal weight gain, blood lead levels, and cesarean section delivery	–
Berkowitz et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L) by maternal PON1 activity (tertile)	76 < 11.0 μg/L, low PON1 62 < 11.0 μg/L, medium PON1 71 < 11.0 μg/L, high PON1 47 > 11.0 μg/L, low PON1 57 > 11.0 μg/L, medium PON1 55 > 11.0 μg/L, high PON1	Mean ± SD = 3,237 ± 456 Mean ± SD = 3,255 ± 436 Mean ± SD = 3,337 ± 444 P-trend > 0.05 Mean ± SD = 3,278 ± 395 Mean ± SD = 3,327 ± 406 Mean ± SD = 3,270 ± 409 P-trend > 0.05	"	Results for TCPy not reported by infant PON1 activity or maternal or infant *PON1* genotype
Berkowitz et al. (2004)	Birth length (cm)	Maternal prenatal urinary TCPy (μg/L)	216 < 11.0 μg/L (limit of detection) 171 > 11.0 μg/L	Mean ± SD = 50.4 ± 2.4 Mean ± SD = 50.8 ± 2.4 $P > 0.05$	"	–
Berkowitz et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L) by maternal PON1 activity (tertile)	75 < 11.0 μg/L, low PON1 62 < 11.0 μg/L, medium PON1 71 < 11.0 μg/L, high PON1 46 > 11.0 μg/L, low PON1 57 > 11.0 μg/L, medium PON1 55 > 11.0 μg/L, high PON1	Mean ± SD = 50.3 ± 2.3 Mean ± SD = 50.1 ± 2.2 Mean ± SD = 50.3 ± 2.3 P-trend > 0.05 Mean ± SD = 50.9 ± 2.3 Mean ± SD = 51.0 ± 2.3 Mean ± SD = 50.8 ± 2.4 P-trend > 0.05	"	–

Berkowitz et al. (2004)	Head circumference (cm)	Maternal prenatal urinary TCPy (μg/L)	216 < 11.0 μg/L (limit of detection) 171 > 11.0 μg/L	Mean ± SD = 33.8 ± 1.7 Mean ± SD = 33.8 ± 1.7 $P > 0.05$	"	–
Berkowitz et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L) by maternal PON1 activity (tertile)	76 < 11.0 μg/L, low PON1 62 < 11.0 μg/L, medium PON1 70 < 11.0 μg/L, high PON1 47 > 11.0 μg/L, low PON1 57 > 11.0 μg/L, medium PON1 55 > 11.0 μg/L, high PON1	Mean ± SD = 33.6 ± 1.8 Mean ± SD = 33.7 ± 1.7 Mean ± SD = 34.1 ± 1.7 P-trend > 0.05 Mean ± SD = 33.3 ± 1.5 Mean ± SD = 34.0 ± 1.5 Mean ± SD = 34.1 ± 1.6 P-trend = 0.014	No difference after additional adjustment for birth weight or birth length, stratification by race/ethnicity, or excluding preterm births	Test for interaction among TCPy level, PON1 activity, and head circumference was not statistically significant ($P > 0.05$)
Berkowitz et al. (2004)	Gestational age (weeks)	Maternal prenatal urinary TCPy (μg/L)	216 < 11.0 μg/L (limit of detection) 171 > 11.0 μg/L	Mean ± SD = 39.3 ± 1.8 Mean ± SD = 39.3 ± 1.7 $P > 0.05$	Race/ethnicity and infant sex No difference after additional adjustment for active and passive cigarette smoking, pre-pregnancy body mass index, maternal weight gain, blood lead levels, and cesarean section delivery	–
Wolff et al. (2007)	Birth weight (g)	Maternal prenatal urinary DAPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = − 25 ± 34, $P = 0.47$ (not creatinine-adjusted) Beta ± SE = − 27 ± 34, $P = 0.43$ (creatinine-adjusted)	Race/ethnicity, maternal PON1 activity, infant sex, and gestational age	Value of 0.5 was added to urinary DAP before log-transformation; 25 samples with < 20 mg/dL creatinine were excluded
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	327	Beta ± SE = − 1.9 ± 29, $P = 0.95$ (not creatinine-adjusted) Beta ± SE = − 2.7 ± 29, $P = 0.92$(creatinine-adjusted)		–
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = − 52 ± 32, $P = 0.099$ (not creatinine-adjusted) Beta ± SE = − 56 ± 32, $P = 0.082$ (creatinine-adjusted)	"	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs ≥ vs. < median by maternal PON1 activity (tertile)	60 DEPs < median, low PON1 53 DEPs < median, medium PON1 45 DEPs < median, high PON1 53 DEPs ≥ median, low PON1 51 DEPs ≥ median, medium PON1 56 DEPs ≥ median, high PON1	Mean ± SE = 3305 ± 53 Mean ± SE = 3348 ± 57 Mean ± SE = 3396 ± 64 Mean ± SE = 3233 ± 56, $P = 0.323$ within PON1 Mean ± SE = 3282 ± 57, $P = 0.392$ within PON1 Mean ± SE = 3279 ± 54, $P = 0.138$ within PON1 P for interaction term in model = 0.878 $P = 0.042$ for high PON1/low DEP vs. low PON1/high DEP	Infant race, sex, gestational age, and creatinine level	Lowest PON1 tertile = slow; highest PON1 tertile = fast $PON1_{192}$ RR = slow, $PON1_{192}$ QQ = fast
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs ≥ vs. < median by maternal $PON1_{192}$ genotype	39 DEPs < median, $PON1_{192}$ RR 84 DEPs < median, $PON1_{192}$ RQ 33 DEPs < median, $PON1_{192}$ QQ 55 DEPs ≥ median, $PON1_{192}$ RR 66 DEPs ≥ median, $PON1_{192}$ RQ 42 DEPs ≥ median, $PON1_{192}$ QQ	Mean ± SE = 3346 ± 69 Mean ± SE = 3278 ± 46 Mean ± SE = 3453 ± 60 Mean ± SE = 3254 ± 63, $P = 0.291$ within PON1192 Mean ± SE = 3285 ± 50, $P = 0.907$ within PON1192 Mean ± SE = 3232 ± 52, $P = 0.005$ within PON1192 P for interaction term in model = 0.0755 $P = 0.020$ for PON1192 QQ/low DEP vs. PON1192 RR/high DEP	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary MDA > 0.3 vs. < 0.3 μg/L (limit of detection)	330	Beta ± SE = 39 ± 52, $P = 0.46$ (not creatinine-adjusted) Beta ± SE = 59 ± 53, $P = 0.27$ (creatinine-adjusted)	Race/ethnicity, maternal PON1 activity, infant sex, and gestational age	–
Wolff et al. (2007)	Birth length (cm)	Maternal prenatal urinary DAPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.13 ± 19, $P = 0.49$ (not creatinine-adjusted) Beta ± SE = −0.13 ± 19, $P = 0.49$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	327	Beta ± SE = −0.12 ± 0.16, $P = 0.44$ (not creatinine-adjusted) Beta ± SE = −0.12 ± 0.16, $P = 0.44$ (creatinine-adjusted)	"	–

Wolff et al. (2007)	"	Maternal prenatal urinary DMPs ≥ vs. < median by maternal PON1 activity (tertile)	60 DMPs < median, low PON1 53 DMPs < median, medium PON1 45 DMPs < median, high PON1 53 DMPs ≥ median, low PON1 51 DMPs ≥ median, medium PON1 56 DMPs ≥ median, high PON1	Mean ± SE = 51.1 ± 0.3 Mean ± SE = 50.3 ± 0.3 Mean ± SE = 50.4 ± 0.3 Mean ± SE = 50.2 ± 0.3, $P = 0.032$ within PON1 Mean ± SE = 50.7 ± 0.3, $P = 0.258$ within PON1 Mean ± SE = 50.8 ± 0.3, $P = 0.418$ within PON1 P for interaction term in model = 0.036 $P = 0.549$ for high PON1/low DMP vs. low PON1/high DMP	Infant race, sex, gestational age, and creatinine level	–
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs ≥ vs. < median by maternal $PON1_{192}$ genotype	39 DMPs < median, $PON1_{192}$ RR 84 DMPs < median, $PON1_{192}$ RQ 33 DMPs < median, $PON1_{192}$ QQ 55 DMPs ≥ median, $PON1_{192}$ RR 66 DMPs ≥ median, $PON1_{192}$ RQ 42 DMPs ≥ median, $PON1_{192}$ QQ	Mean ± SE = 50.6 ± 0.4 Mean ± SE = 50.4 ± 0.3 Mean ± SE = 51.0 ± 0.3 Mean ± SE = 49.9 ± 0.3, $P = 0.164$ within PON1192 Mean ± SE = 50.7 ± 0.3, $P = 0.158$ within PON1192 Mean ± SE = 50.8 ± 0.3, $P = 0.695$ within PON1192 P for interaction term in model = 0.230 $P = 0.019$ for PON1192 QQ/low DMP vs. PON1192 RR/high DMP	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.02 ± 0.18, $P = 0.93$ (not creatinine-adjusted) Beta ± SE = 0.017 ± 0.18, $P = 0.924$ (creatinine-adjusted)	Race/ethnicity, maternal PON1 activity, infant sex, and gestational age	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Wolff et al. (2007)	"	Maternal prenatal urinary MDA > 0.3 vs. < 0.3 µg/L (limit of detection)	330	Beta ± SE = −0.16 ± 0.28, $P = 0.56$ (not creatinine-adjusted) Beta ± SE = −0.032 ± 30, $P = 0.91$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	Ponderal index (g/cm^3)	Maternal prenatal urinary DAPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.002 ± 0.023, $P = 0.93$ (not creatinine-adjusted) Beta ± SE = −0.003 ± 0.023, $P = 0.91$ (creatinine-adjusted)	"	No significant interactions between DAPs and PON1 were detected for ponderal index (results NR)
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	327	Beta ± SE = 0.01 ± 0.02, $P = 0.47$ (not creatinine-adjusted) Beta ± SE = 0.01 ± 0.02, $P = 0.48$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.04 ± 0.02, $P = 0.087$ (not creatinine-adjusted) Beta ± SE = −0.04 ± 0.02, $P = 0.077$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary MDA > 0.3 vs. < 0.3 µg/L (limit of detection)	330	Beta ± SE = 0.039 ± 0.035, $P = 0.27$ (not creatinine-adjusted) Beta ± SE = 0.035 ± 0.036, $P = 0.33$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	Head circumference (cm)	Maternal prenatal urinary DAPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.26 ± 0.13, $P = 0.045$ (not creatinine-adjusted) Beta ± SE = −0.25 ± 0.13, $P = 0.056$ (creatinine-adjusted)	"	No significant interactions between DAPs and PON1 were detected for head circumference (results NR)
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	327	Beta ± SE = −0.16 ± 0.11, $P = 0.14$ (not creatinine-adjusted) Beta ± SE = −0.15 ± 0.11, $P = 0.16$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = −0.067 ± 0.12, $P = 0.57$ (not creatinine-adjusted) Beta ± SE = −0.052 ± 0.12, $P = 0.67$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary MDA > 0.3 vs. < 0.3 µg/L (limit of detection)	330	Beta ± SE = 0.15 ± 0.19, $P = 0.44$ (not creatinine-adjusted) Beta ± SE = 0.23 ± 0.20, $P = 0.25$ (creatinine-adjusted)	"	–

Wolff et al. (2007)	Gestational age (weeks)	Maternal prenatal urinary DAPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = 0.03 ± 0.14, $P = 0.81$ (not creatinine-adjusted) Beta ± SE = 0.03 ± 0.14, $P = 0.83$ (creatinine-adjusted)	Race/ethnicity, maternal PON1 activity, and infant sex	–
Wolff et al. (2007)	"	Maternal prenatal urinary DMPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	327	Beta ± SE = − 0.029 ± 0.12, $P = 0.80$ (not creatinine-adjusted) Beta ± SE = − 0.030 ± 0.12, $P = 0.80$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary DEPs (nmol/L or nmol/g creatinine, $\log_{10}$ scale)	318	Beta ± SE = − 0.006 ± 0.13, $P = 0.996$ (not creatinine-adjusted) Beta ± SE = − 0.004 ± 0.13, $P = 0.97$ (creatinine-adjusted)	"	–
Wolff et al. (2007)	"	Maternal prenatal urinary MDA > 0.3 vs. < 0.3 µg/L (limit of detection)	330	Beta ± SE = − 0.28 ± 0.21, $P = 0.18$ (not creatinine-adjusted) Beta ± SE = − 0.30 ± 0.22, $P = 0.16$ (creatinine-adjusted)	"	–
Eskenazi et al. (2004)	Length of gestation (weeks)	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	485 with DAPs 485 with DMPs 486 with DEPs	Beta = − 0.20 (− 0.55, 0.15) Beta = − 0.41 (− 0.75, − 0.07) Beta = − 0.16 (− 0.53, 0.22)	Timing of urine collection, timing of entry into prenatal care, maternal age, parity, maternal country of birth, and poverty level	Gestational age based on medical record; results similar when based on maternal self-reported date of last menstrual period Results persisted when metabolite levels were controlled for creatinine Inverse association with DMPs was most apparent for specimens collected after 22 weeks of gestation Associations of DEAMPY, IMPY, CMHC, and CIT with birth outcomes not analyzed due to small percentage of women with detectable levels
Eskenazi et al. (2004)	"	Maternal prenatal urinary MDA (µg/L)	233 undetectable 74 detectable < median 75 detectable ≥ median	Beta = referent Beta = − 0.13 (− 0.55, 0.30) Beta = − 0.21 (− 0.62, 0.20)	"	
Eskenazi et al. (2004)	"	Maternal prenatal urinary TCPy (µg/L)	41 undetectable 220 detectable < median 221 detectable ≥ median	Beta = referent Beta = − 0.17 (− 0.74, 0.40) Beta = − 0.06 (− 0.63, 0.51)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary PNP (µg/L)	124 undetectable 179 detectable < median 179 detectable ≥ median	Beta = referent Beta = − 0.37 (− 0.76, 0.02) Beta = 0.18 (− 0.21, 0.57)	"	–

(Continued)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/ min/mL)	340 maternal blood, prenatal	Beta = 0.01 (−0.15, 0.17)	"	When gestational age was based on maternal self-reported date of last menstrual period, beta for lower cholinesterase in maternal blood = 1.1 days, $P = 0.04$
			357 maternal blood, delivery	Beta = 0.09 (−0.04, 0.23)		
			292 cord blood	Beta = 0.34 (0.13, 0.55)		
Eskenazi et al. (2004)	"	Maternal/cord plasma butyrylcholinesterase (μmol/min/mL)	340 maternal plasma, prenatal	Beta = −0.2 (−0.64, 0.27)	"	–
			357 maternal plasma, delivery	Beta = −0.1 (−0.48, 0.36)		
			292 cord plasma	Beta = −0.2 (−0.78, 0.32)		
Eskenazi et al. (2004)	Birth weight (g)	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	485 with DAPs	Beta = 42 (−46, 131)	Timing of urine collection, timing of entry into prenatal care, maternal age, parity, infant sex, maternal country of birth, pregnancy weight gain, body mass index, poverty level, gestational age, and gestational age squared	–
			485 with DMPs	Beta = 41 (−40, 122)		
			486 with DEPs	Beta = 52 (−40, 144)		
Eskenazi et al. (2004)	"	Maternal prenatal urinary MDA (μg/L)	233 undetectable	Beta = referent	"	–
			74 detectable < median	Beta = −45 (−154, 63)		
			75 detectable ≥ median	Beta = 56 (−49, 161)		
Eskenazi et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L)	41 undetectable	Beta = referent	"	–
			220 detectable < median	Beta = −6 (−138, 126)		
			221 detectable ≥ median	Beta = 27 (−106, 159)		
Eskenazi et al. (2004)	"	Maternal prenatal urinary PNP (μg/L)	124 undetectable	Beta = referent	"	–
			179 detectable < median	Beta = 34 (−57, 125)		
			179 detectable ≥ median	Beta = 49 (−42, 140)		
Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/ min/mL)	340 maternal blood, prenatal	Beta = 8 (−35, 52)	"	–
			357 maternal blood, delivery	Beta = 6 (−30, 43)		
			292 cord blood	Beta = 12 (−46, 70)		
Eskenazi et al. (2004)	"	Maternal/cord plasma butyrylcholinesterase (μmol/min/mL)	340 maternal plasma, prenatal	Beta = 56 (−67, 179)	"	–
			357 maternal plasma, delivery	Beta = −90 (−206, 25)		
			292 cord plasma	Beta = 111 (−35, 257)		

Eskenazi et al. (2004)	Body length (cm)	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	485 with DAPs 485 with DMPs 486 with DEPs	Beta = 0.52 (−0.01, 1.05) Beta = 0.42 (−0.07, 0.91) Beta = 0.40 (−0.15, 0.94)	"	No association when metabolite levels were controlled for creatinine Positive association with DAPs did not vary substantially by week of prenatal urine collection
Eskenazi et al. (2004)	"	Maternal prenatal urinary MDA (μg/L)	233 undetectable 74 detectable < median 75 detectable ≥ median	Beta = referent Beta = −0.53 (−1.18, 0.11) Beta = 0.14 (−0.48, 0.76)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L)	41 undetectable 220 detectable < median 221 detectable ≥ median	Beta = referent Beta = 0.09 (−−0.70, 0.87) Beta = 0.44 (−0.35, 1.22)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary PNP (μg/L)	124 undetectable 179 detectable < median 179 detectable ≥ median	Beta = referent Beta = 0.60 (0.06, 1.13) Beta = 0.41 (−0.13, 0.94)	"	–
Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/min/mL)	340 maternal blood, prenatal 357 maternal blood, delivery 292 cord blood	Beta = 0.05 (−0.20, 0.29) Beta = 0.05 (−0.17, 0.27) Beta = −0.01 (−0.35, 0.34)	"	–
Eskenazi et al. 2004	"	Maternal/cord plasma butyrylcholinesterase (μmol/min/mL)	340 maternal plasma, prenatal 357 maternal plasma, delivery 292 cord plasma	Beta = 0.07 (−0.63, 0.78) Beta = 0.05 (−0.65, 0.75) Beta = 0.23 (−0.65, 1.12)	"	–
Eskenazi et al. (2004)	Head circumference (cm)	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	485 with DAPs 485 with DMPs 486 with DEPs	Beta = 0.32 (0.03, 0.62) Beta = 0.25 (−0.02, 0.52) Beta = 0.28 (−0.02, 0.59)	"	Results persisted when metabolite levels were controlled for creatinine Positive association with DAPs did not vary substantially by week of prenatal urine collection
Eskenazi et al. (2004)	"	Maternal prenatal urinary MDA (μg/L)	233 undetectable 74 detectable < median 75 detectable ≥ median	Beta = referent Beta = −0.16 (−0.52, 0.19) Beta = 0.11 (−0.24, 0.46)	"	–
Eskenazi et al. 2004	"	Maternal prenatal urinary TCPy (μg/L)	41 undetectable 220 detectable < median 221 detectable ≥ median	Beta = referent Beta = 0.06 (−0.37, 0.49) Beta = 0.04 (−0.39, 0.47)	"	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2004)	"	Maternal prenatal urinary PNP (μg/L)	124 undetectable 179 detectable < median 179 detectable ≥ median	Beta = referent Beta = 0.18 (−0.12, 0.48) Beta = 0.29 (−0.01, 0.58)	"	–
Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/min/mL)	340 maternal blood, prenatal 357 maternal blood, delivery 292 cord blood	Beta = 0.06 (−0.09, 0.21) Beta = −0.07 (−0.19, 0.05) Beta = −0.04 (−0.23, 0.14)	"	–
Eskenazi et al. (2004)	"	Maternal/cord plasma butyrylcholinesterase (μmol/min/mL)	340 maternal plasma, prenatal 357 maternal plasma, delivery 292 cord plasma	Beta = 0.12 (−0.31, 0.56) Beta = −0.07 (−0.45, 0.31) Beta = −0.03 (−0.50, 0.45)	"	–
Eskenazi et al. (2004)	Ponderal index (g/cm^3)	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	485 with DAPs 485 with DMPs 486 with DEPs	Beta = −0.04 (−0.12, 0.04) Beta = −0.03 (−0.10, 0.04) Beta = −0.01 (−0.09, 0.07)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary MDA (μg/L)	233 undetectable 74 detectable < median 75 detectable ≥ median	Beta = referent Beta = 0.05 (−0.05, 0.14) Beta = 0.02 (−0.07, 0.12)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary TCPy (μg/L)	41 undetectable 220 detectable < median 221 detectable ≥ median	Beta = referent Beta = −0.01 (−0.12, 0.11) Beta = −0.04 (−0.16, 0.08)	"	–
Eskenazi et al. (2004)	"	Maternal prenatal urinary PNP (μg/L)	124 undetectable 179 detectable < median 179 detectable ≥ median	Beta = referent Beta = −0.08 (−0.16, 0.0) Beta = −0.03 (−0.11, 0.05)	"	–
Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/min/mL)	340 maternal blood, prenatal 357 maternal blood, delivery 292 cord blood	Beta = 0.00 (−0.03, 0.03) Beta = 0.00 (−0.03, 0.03) Beta = 0.02 (−0.03, 0.07)	"	–
Eskenazi et al. (2004)	"	Maternal/cord plasma butyrylcholinesterase (μmol/min/mL)	340 maternal plasma, prenatal 357 maternal plasma, delivery 292 cord plasma	Beta = 0.03 (−0.06, 0.12) Beta = −0.07 (−0.16, 0.03) Beta = 0.05 (−0.07, 0.17)	"	–
Eskenazi et al. (2004)	Preterm delivery	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	32 (6.6%) preterm	"not associated" (results NR)	NR	Preterm delivery: birth at < 37 completed weeks of gestation

Eskenazi et al. (2004)	"	Maternal/cord blood cholinesterase (μmol/min/mL; per unit decrease)	NR preterm with maternal prenatal blood	Odds ratio = 1.6 (1.0, 2.5)	NR	–
			NR preterm with cord blood	Odds ratio = 2.3 (1.1, 4.8)		
Eskenazi et al. (2004)	Low birth weight	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	18 (3.7%) low birth weight	"not associated" (results NR)	NR	Low birth weight: < 2,500 g
Eskenazi et al. (2004)	"	Cord blood cholinesterase (μmol/min/mL; per unit decrease)	11 low birth weight with cord blood	Odds ratio = 4.3 (1.1, 17.5)	NR	Six of 11 infants with low birth weight were also preterm
Eskenazi et al. (2004)	Small for gestational age birth	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	23 (48%) small for gestational age birth	"not associated" (results NR)	NR	Small for gestational age birth: birth weight < 10th percentile for gestational age according to ethnicity, parity, and infant sex
Harley et al. (2011)	Gestational age (weeks)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = −0.8 (−2.0, 0.2) Beta = −0.3 (−0.8, 0.3) Beta = 0.1 (−0.7, 0.8) P-interaction = 0.36 Beta = −1.0 (−2.0, 0.0) Beta = −0.2 (−0.7, 0.3) Beta = 0.2 (−0.7, 1.1) P-interaction = 0.21	Timing of urine collection, timing of entry into prenatal care, maternal age, parity, maternal country of birth, and household income	Arylesterase activity (a marker of PON1 enzyme quantity) was lowest in mothers and infants with $PON1_{-108}$ TT (but highest in mothers with $PON1_{192}$ QQ) *P*araoxonase activity (a marker of PON1 enzyme activity) was lowest in mothers and infants with either $PON1_{-108}$ TT or $PON1_{192}$ QQ, especially both
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = −0.6 (−1.6, 0.4) Beta = −0.3 (−0.8, 0.2) Beta = 0.0 (−0.7, 0.7) P-interaction = 0.49 Beta = −0.7 (−1.7, 0.2) Beta = −0.3 (−0.7, 0.1) Beta = 0.3 (−0.5, 1.2) P-interaction = 0.25	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = −1.0 (−2.1, 0.1) Beta = −0.2 (−0.8, 0.3) Beta = 0.6 (−0.2, 1.4) P-interaction = 0.09 Beta = −1.0 (−2.1, 0.2) Beta = 0.1 (−0.4, 0.6) Beta = −0.3 (−1.2, 0.6) P-interaction = 0.17	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity: 108 tertile 1 108 tertile 2 108 tertile 3	 Beta = 0.3 (−0.5, 1.2) Beta = −0.4 (−1.2, 0.4) Beta = −0.2 (−0.9, 0.5) P-interaction = 0.17	"	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity:		"	–
			108 tertile 1	Beta = 0.5 (−0.3, 1.2)		
			108 tertile 2	Beta = −0.4 (−1.1, 0.4)		
			108 tertile 3	Beta = −0.4 (−1.0, 0.3)		
				P-interaction = 0.16		
Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity:		"	–
			108 tertile 1	Beta = −0. (−1.2, 0.4)		
			108 tertile 2	Beta = −0.7 (−1.5, 0.1)		
			108 tertile 3	Beta = 0.5 (−0.3, 1.2)		
				P-interaction = 0.69		
Harley et al. (2011)	Birth weight (g)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT	Beta = −131.3 (−393.3, 130.8)	Timing of urine collection, timing of entry into prenatal care, maternal age, parity, maternal country of birth, household income, pre-pregnancy body mass index, maternal weight gain during pregnancy, infant sex, gestational age, and gestational age squared	–
			225 $PON1_{-108}$ CT	Beta = 147.2 (18.5, 275.7)		
			131 $PON1_{-108}$ CC	Beta = 22.1 (−182.0, 226.3)		
				P-interaction = 0.06		
			108 $PON1_{192}$ QQ	Beta = −60.2 (−266.3, 145.9)		
			222 $PON1_{192}$ QR	Beta = 79.4 (−48.5, 207.3)		
			106 $PON1_{192}$ RR	Beta = 142.3 (−114.6, 399.3)		
				P-interaction = 0.20		
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT	Beta = −135.2 (−373.8, 103.3)	"	–
			225 $PON1_{-108}$ CT	Beta = 134.6 (14.9, 254.2)		
			131 $PON1_{-108}$ CC	Beta = 46.8 (−132.5, 226.1)		
				P-interaction = 0.05		
			108 $PON1_{192}$ QQ	Beta = −80.6 (−269.2, 107.9)		
			222 $PON1_{192}$ QR	Beta = 89.9 (−27.8, 207.5)		
			106 $PON1_{192}$ RR	Beta = 72.9 (−169.2, 315.0)		
				P-interaction = 0.16		
Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT	Beta = −55.3 (−320.5, 209.8)	"	–
			225 $PON1_{-108}$ CT	Beta = 120.8 (−14.8, 256.4)		
			131 $PON1_{-108}$ CC	Beta = 45.4 (−174.6, 265.4)		
				P-interaction = 0.35		
			108 $PON1_{192}$ QQ	Beta = 20.5 (−210.5, 251.5)		
			222 $PON1_{192}$ QR	Beta = 67.2 (−63.3, 197.5)		
			106 $PON1_{192}$ RR	Beta = 258.8 (23.9, 493.6)		
				P-interaction = 0.30		
Harley et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity:		"	–
			108 tertile 1	Beta = −14.6 (−263.4, 234.3)		
			108 tertile 2	Beta = 63.8 (−131.1, 258.7)		
			108 tertile 3	Beta = 92.2 (−106.6, 291.1)		
				P-interaction = 0.39		
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity:		"	–
			108 tertile 1	Beta = 14.8 (−216.2, 245.7)		
			108 tertile 2	Beta = 83.9 (−94.8, 262.6)		
			108 tertile 3	Beta = 60.2 (−120.2, 240.7)		
				P-interaction = 0.29		

Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity: 108 tertile 1 108 tertile 2 108 tertile 3	 Beta = −55.2 (−295.6, 185.2) Beta = 48.7 (−134.4, 231.8) Beta = 231.4 (19.1, 443.6) P-interaction = 0.31	"	–
Harley et al. (2011)	Head circumference (cm)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = 0.1 (−0.6, 0.9) Beta = 0.2 (−0.2, 0.6) Beta = 0.6 (−0.1, 1.3) P-interaction = 0.08 Beta = −0.3 (−1.0, 0.4) Beta = 0.2 (−0.2, 0.6) Beta = 0.7 (−0.1, 1.5) P-interaction = 0.01	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = 0.2 (−0.5, 0.8) Beta = 0.1 (−0.3, 0.5) Beta = 0.5 (−0.2, 1.1) P-interaction = 0.12 Beta = −0.4 (−1.0, 0.2) Beta = 0.2 (−0.2, 0.5) Beta = 0.5 (−0.3, 1.3) P-interaction = 0.01	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by child genotype	76 $PON1_{-108}$ TT 225 $PON1_{-108}$ CT 131 $PON1_{-108}$ CC 108 $PON1_{192}$ QQ 222 $PON1_{192}$ QR 106 $PON1_{192}$ RR	Beta = −0.1 (−0.8, 0.7) Beta = 0.2 (−0.2, 0.6) Beta = 0.6 (−0.2, 1.4) P-interaction = 0.19 Beta = 0.1 (−0.7, 0.9) Beta = 0.1 (−0.3, 0.5) Beta = 0.7 (0.0, 1.5) P-interaction = 0.27	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity: 108 tertile 1 108 tertile 2 108 tertile 3	 Beta = −0.2 (−1.0, 0.5) Beta = 0.3 (−0.3, 0.9) Beta = 0.8 (0.1, 1.4) P-interaction = 0.32	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity: 108 tertile 1 108 tertile 2 108 tertile 3	 Beta = −0.1 (−0.8, 0.5) Beta = 0.1 (−0.4, 0.7) Beta = 0.7 (0.1, 1.3) P-interaction = 0.36	"	–
Harley et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity	PON1 quantity: 108 tertile 1 108 tertile 2 108 tertile 3	 Beta = −0.5 (−1.2, 0.2) Beta = 0.4 (−0.2, 1.0) Beta = 0.7 (−0.1, 1.4) P-interaction = 0.48	"	–
Barr et al. (2010)	Birth weight (g)	Maternal pre/perinatal or cord serum chlorpyrifos (ng/g) > 75th vs. ≤ 75th percentile	138 maternal serum 148 cord serum	Mean ± SD = 3053 ± 111 vs. 3548 ± 448, $P = 0.268$ Mean ± SD = 3581 ± 422 vs. 3544 ± 433, $P = 0.408$	Maternal age, primigravida, race, pre-pregnancy body mass index, infant sex, and gestational age	Results were similar when pesticide levels were dichotomized at the 90th percentile (results NR)
Barr et al. (2010)	Head circumference (cm)	"	138 maternal serum 148 cord serum	Mean ± SD = 33.4 ± 0.6 vs. 35.0 ± 1.3, $P = 0.229$ Mean ± SD = 34.1 ± 1.4 vs. 35.0 ± 1.2, $P = 0.989$	"	

(Continued)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Barr et al. (2010)	Abdominal circumference (in)	"	138 maternal serum	Mean ± SD = 29.2 ± 0.8 vs. 32.0 ± 2.7, $P = 0.201$	"	
			148 cord serum	Mean ± SD = 32.5 ± 2.3 vs. 32.0 ± 2.7, $P = 0.346$		
Barr et al. (2010)	Birth length (cm)	"	138 maternal serum	Mean ± SD = 49.8 ± 0.2 vs. 51.3 ± 3.0, $P = 0.686$	"	
			148 cord serum	Mean ± SD = 50.9 ± 1.7 vs. 51.4 ± 3.1, $P = 0.318$		
Wang et al. (2011)	Length of gestation (weeks)	Maternal perinatal urinary DAPs (log scale) [unit (nmol/L or nmol/g creatinine) and log base not specified]	187	Dimethylphosphate beta = − 0.05 (− 0.52–0.33) Dimethylthiophosphate beta = 0.15 (− 1.21–1.03) Diethylphosphate beta = 0.11 (− 1.27–0.52) Diethylthiophosphate beta = 0.13 (− 0.92–0.64) Diethyldithiophosphate beta = − 0.03 (− 0.04–0.06) DAPs beta = 0.04 (− 0.35–0.59)	Maternal height, pregnancy weight gain, and family income	Some apparent reporting errors (e.g., missing "−" signs) are corrected here based on reported P-values Results were unchanged when preterm infants were excluded (results NR)
Wang et al. (2011)	"	"	91 infant girls	Dimethylphosphate beta = 0.39 (− 0.13, 0.63) Dimethylthiophosphate beta = 0.31 (− 0.08–0.63) Diethylphosphate beta = − 1.79 (− 2.82 to − 0.76) [Boys: diethylphosphate beta = 0.17, $P = 0.164$] Diethylthiophosphate beta = 0.72 (− 0.28, 1.16) Diethyldithiophosphate beta = 0.09 (− 0.35–0.53) DAPs beta = − 0.03 (− 0.81–0.61)	"	–
Wang et al. (2011)	Birth weight (g)	"	187	Dimethylphosphate beta = − 18 (− 151–109) Dimethylthiophosphate beta = 84 (− 50–304) Diethylphosphate beta = 135 (− 143–402) Diethylthiophosphate beta = − 112 (− 318–159) Diethyldithiophosphate beta = 4 (− 161–313) DAPs beta = 69 (− 74–212)	Gestational age, maternal height, pregnancy weight gain, and family income	–

Wang et al. (2011)	"	"	91 infant girls	Dimethylphosphate beta = − 48 (− 192–218) Dimethylthiophosphate beta = 166 (− 40–473) Diethylphosphate beta = 174 (− 287–529) Diethylthiophosphate beta = − 272 (− 499–208) Diethyldithiophosphate beta = 45 (− 278–412) DAPs beta = − 6 (− 286–240)	"	–
Wang et al. (2011)	Body length (cm)	"	187	Dimethylphosphate beta = − 0.01 (− 0.67–0.61) Dimethylthiophosphate beta = − 0.04 (− 0.67–1.07) Diethylphosphate beta = 0.12 (− 0.65–2.02) Diethylthiophosphate beta = − 0.16 (− 2.03–0.31) Diethyldithiophosphate beta = − 0.01 (− 1.22–1.10) DAPs beta = 0.03 (− 0.47–0.73)	Gestational age, maternal height, pregnancy weight gain, and family income	–
Wang et al. (2011)	"	"	91 infant girls	Dimethylphosphate beta = − 0.06 (− 0.73–0.48) Dimethylthiophosphate beta = 0.24 (− 0.11–1.40) Diethylphosphate beta = 0.33 (− 1.00–1.41) Diethylthiophosphate beta = − 0.16 (− 1.54–0.55) Diethyldithiophosphate beta = − 0.09 (− 1.40–0.64) DAPs beta = − 0.17(− 1.84–0.21)	"	–

(*Continued*)

Table 2. (Continued)

Reference	Outcome	Exposure	Number of subjects/ events	Estimate of association (95% CI)	Adjustment factors	Comments
Rauch et al. (2012)	Length of gestation (weeks)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −0.5 (−0.8, −0.1) Beta = −0.1 (−0.9, 0.6) Beta = −0.7 (−1.1, −0.3) P-interaction by race = 0.10	Mother's age, mother's race (unless stratified), household income, marital status, parity category, $\log_{10}$-transformed blood lead, and $\log_{10}$-transformed cotinine	Results were similar when excluding mothers with abruptio placenta, placenta previa, chorioamnionitis, pre-eclampsia, or pregnancy-induced hypertension, or when excluding mothers with urinary creatinine < 20 mg/dL Results were similar or attenuated (beta for DAPs and gestational age = −0.2 [−0.4, 0.0]; beta for DAP and birth weight = −0.88 [−213, 37]) when restricted to full-term births Results did not vary significantly by child sex (P-interaction > 0.3) Results were "modestly attenuated" when based on non-creatinine-adjusted DAPs (beta for gestational age = −0.3 [−0.7, 0.0]; beta for birth weight = −100 [−232, 32]), and "slightly attenuated" when based on individual urine specimens from 16 or 26 weeks of gestation
Rauch et al. (2012)	"	"	55 *PON1*$_{192}$ RR 107 *PON1*$_{192}$ QR 111 *PON1*$_{192}$ QQ	Beta = −0.3 (−1.2, 0.5) Beta = −0.9 (−1.6, −0.3) Beta = −0.5 (−1.1, 0.0) P− interaction by genotype = 0.04 for QR, 0.09 for QQ	"	When stratified by race, effect estimates were also stronger in heterozygous groups
			118 *PON1*$_{-108}$ CC 106 *PON1*$_{-108}$ CT 46 *PON1*$_{-108}$ TT	Beta = −0.3 (−0.9, 0.3) Beta = −1.0 (−1.6, −0.4) Beta = −0.1 (−1.0, 0.8) P− interaction by genotype = 0.04 for CT, 0.31 for TT		
Rauch et al. (2012)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −0.4 (−0.7, 0.0) Beta = 0.0 (−0.7, 0.6) Beta = −0.6 (−0.9, −0.2) P-interaction by race = 0.09	"	–
Rauch et al. (2012)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale, creatinine-standardized	306 total 93 black mothers 213 white mothers	Beta = −0.2 (−0.5, 0.1) Beta = −0.1 (−0.8, 0.5) Beta = −0.3 (−0.7, 0.0) P-interaction by race = 0.47	"	–
Rauch et al. (2012)	Birth weight (g)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −151 (−287, −16) Beta = −188 (−395, 19) Beta = −118 (−296, 60) P-interaction by race = 0.46	"	–

Rauch et al. (2012)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −124 (−245, −2) Beta = −142 (−333, 50) Beta = −96 (−254, 62) P-interaction by race = 0.46	"	–
Rauch et al. (2012)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale, creatinine-standardized	306 total 93 black mothers 213 white mothers	Beta = −65 (−180, 51) Beta = −162 (−340, 16) Beta = −39 (−189, 111) P-interaction by race = 0.39	"	–
Rauch et al. (2012)	Birth weight, adjusted for gestational age (g)	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −40 (−146, 65) Beta = −158 (−297, −18) Beta = 60 (−84, 204) P-interaction by race = 0.02	"	–
Rauch et al. (2012)	"	"	55 *PON1*$_{192}$ RR 107 *PON1*$_{192}$ QR 111 *PON1*$_{192}$ QQ	Beta = −71 (−384, 242) Beta = −454 (−707, −201) Beta = −2 (−231, 228) P− interaction by genotype = 0.02 for QR, 0.76 for QQ vs. RR	"	When stratified by race, effect estimates were also stronger in heterozygous groups
			118 *PON1*$_{-108}$ CC 106 *PON1*$_{-108}$ CT 46 *PON1*$_{-108}$ TT	Beta = −119 (−340, 103) Beta = −299 (−520, −78) Beta = 85 (−361, 530) P− interaction by genotype = 0.15 for CT, 0.12 for TT vs. CC		
Rauch et al. (2012)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale, creatinine-standardized)	306 total 93 black mothers 213 white mothers	Beta = −38 (−133, 56) Beta = −139 (−267, −10) Beta = 49 (−78, 177) P-interaction by race = 0.02	"	–
Rauch et al. (2012)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale, creatinine-standardized	306 total 93 black mothers 213 white mothers	Beta = −9 (−99, 80) Beta = −131 (−251, −11) Beta = 41 (−78, 160) P-interaction by race = 0.08	"	–
Wickerham et al. (2012)	Birth weight (g)	Number of organophosphate pesticides (of 8 tested) detected in cord serum	112	Beta = 6.59 (−210, 222)	Gestational age, maternal age, maternal body mass index at early pregnancy, and maternal hemoglobin at delivery	Of 20 pesticides measured, the mean ± SD number detected in cord serum per subject was 4.6 ± 1.9, with a maximum of 10; 98.3% had at least one pesticide detected
Wickerham et al. (2012)	"	Chlorpyrifos, diazinon, fonofos, malathion, parathion-ethyl, parathion-methyl, profenofos, or terbufos in cord serum, detectable vs. non-detectable or 3-level ordinal variables	"	No significant associations (results NR)	"	–

DAP dialkyl phosphate, *DEP* diethyl phosphate, *DMP* dimethyl phosphate, *MDA* malathion dicarboxylic acid, *NR* not reported, *PNP* 4-nitrophenol, *PON1* paraoxonase 1, *SD* standard deviation, *SE* standard error, *TCPy* 3,5,6-trichloro-2-pyridinol.

404 consecutive healthy, primiparous pregnant women with a singleton pregnancy at ≤ 26 weeks of gestation in 1998–2001 (Table 1) (Berkowitz et al. 2004, Wolff et al. 2007). OP metabolites, including TCPy (a metabolite of chlorpyrifos and chlorpyrifos methyl), MDA (a metabolite of malathion), and six DAP metabolites were measured in maternal urine collected during the third trimester. Median concentrations were 7.6 µg/L (below the limit of detection [LOD] of 11.0 µg/L) (interquartile range [IQR] = 1.6–32.5) for TCPy (Berkowitz et al. 2004), < 0.3 µg/L (LOD) (range ≤ 0.3–15.8) for MDA, 75.9 nmol/L (range = < 1–4 [LOD]–4987) for DAPs, 42.2 nmol/L (range ≤ 1–4–4903) for DMPs, and 18.8 nmol/L (range ≤ 1–4–429) for DEPs (Wolff et al. 2007). In addition, five genetic polymorphisms (*Q192R*, *L55M*, *C–909 G*, *A–162 G*, and *C–108 T*) in the *PON1* gene, PON1 enzymatic activity against phenyl acetate, and butyrylcholinesterase (BChE) enzymatic activity against butyrylthiocholine were assessed in the third-trimester maternal blood and umbilical cord blood. PON1 acts as a detoxifying enzyme for OP metabolites, and higher-activity alleles (e.g., $PON1_{192}$ QQ and $PON1_{-108}$ CC) and higher enzyme levels are hypothesized to protect against potential adverse health effects of OP exposure. A recent paper has found that PON1 activity with phenyl acetate as a substrate may not be a reliable index of the quantity of PON1 protein, because the hydrolysis of phenyl acetate is not independent of genotype (McDaniel et al. 2014).

When maternal prenatal urinary TCPy concentration was dichotomized at the LOD (11.0 µg/L), no significant association was observed with birth weight, birth length, or head circumference after multivariate adjustment, based on 387 subjects (Table 2) (Berkowitz et al. 2004). Moreover, none of the three birth outcomes differed significantly by the presence of maternal prenatal urinary TCPy within strata of low, medium, or high maternal PON1 activity. Log_{10}-transformed concentration of maternal prenatal urinary DAPs, with or without creatinine adjustment, was not significantly associated with birth weight, birth length, ponderal index, or gestational age, but it was significantly inversely associated with head circumference (beta = − 0.26 cm, standard error [SE] = 0.13, $P = 0.045$ without creatinine adjustment; similar results with creatinine adjustment) (Wolff et al. 2007). Log_{10}-transformed concentrations of maternal prenatal urinary DMPs and DEPs also were not significantly associated with birth weight, birth length, ponderal index, head circumference, or gestational age, nor was maternal prenatal urinary MDA concentration, dichotomized at the LOD (0.3 µg/L), significantly associated with any of these outcomes. There was some evidence that maternal prenatal urinary DEPs interacted with maternal PON1 activity and $PON1_{192}$ genotype. Birth weights were significantly lower among those with higher DEP concentrations and lower PON1 activity or the $PON1_{192}$ RR (low-activity) genotype, compared with those with lower DEP concentrations and higher PON1 activity ($P = 0.042$) or the $PON1_{192}$ QQ (high-activity) genotype ($P = 0.020$). Also, birth length was significantly shorter among those with higher maternal prenatal urinary DMP concentrations than among those with lower DMP concentrations within the stratum of lower maternal PON1 activity ($P = 0.032$), and among those with higher maternal prenatal urinary DMP concentrations and the $PON1_{192}$ RR genotype, compared with those with lower prenatal DMP and the $PON1_{192}$ QQ genotype ($P = 0.019$). However, such interactions apparently were not detected (or at least were not reported) for total DAPs or for other birth outcomes.

Notable strengths of the Mount Sinai CECS are comparable to those of the CCCEH study, and include the personal measurement of OP metabolites, the detailed characterization of the study cohort, and the prospective design, with exposures measured during the third trimester of pregnancy. Participants were enrolled during a narrower time window, preventing evaluation of associations before and after 2001, but information on *PON1* genotypes and PON1 and BChE activity enabled the assessment of putatively susceptible subgroups. Several limitations are also shared between the cohorts, including the collection of single biospecimens for exposure assessment, the small number of subjects and multiple hypothesis testing in stratified analyses, and possible modest confounding. Thirty-three percent of eligible women consented to participate, and 74% of participants were ultimately included in the analysis after exclusions due to medical issues, lack of biospecimens, change of hospital or residence, or refusal, with shorter follow-up for younger and less-educated mothers. If participation and/or follow-up were associated with both OP exposure and subsequent birth outcomes, then selection bias could have distorted the results in an unpredictable direction. Maternal prenatal urinary TCPy and MDA levels were analyzed as dichotomous variables (detectable or non-detectable), precluding analyses of exposure–response trends. As described earlier in this paper, urinary DAP concentrations are unlikely to accurately reflect long-term exposure to OP insecticides. They also do not enable identification of associations with specific OP insecticides, which differ substantially in acute toxicity.

Taken together, the results in the Mount Sinai CECS cohort show no association between maternal prenatal levels of TCPy or MDA and birth outcomes. Higher maternal prenatal DAP levels were associated with smaller head circumference, but not birth weight, birth length, ponderal index, or gestational age, while maternal prenatal DMP and DEP levels were not significantly associated with any of these outcomes in the combined cohort. The association with head circumference alone could be interpreted as indicative of a neurotoxic effect or, alternatively, a chance for finding amid predominantly null results. Some interactions in the expected direction (assuming greater susceptibility in those with lower PON1 activity) were detected between prenatal DMPs or DEPs and maternal PON1 activity or $PON1_{192}$ genotype in association with birth weight or birth length. However, no explanation was provided for why DMPs would interact with PON1 activity and genotype in relation to birth length but not weight or other birth outcomes, whereas DEPs would interact with PON1 activity and genotype in relation to birth weight but not length or other birth outcomes. The internally inconsistent findings indicate no clear relationship between prenatal exposure to OP insecticides and fetal growth or gestational age.

Center for the Health Assessment of Mothers and Children of Salinas

The CHAMACOS prospectively enrolled pregnant women entering prenatal care at < 20 weeks of gestation between 1999 and 2000 in a primarily Latino, low-income, farm-

worker population in the Salinas Valley of California (Table 1) (Eskenazi et al. 2004, Harley et al. 2011). Six DAP metabolites were measured in maternal spot urine specimens collected at approximately 13–14 weeks and 26–27 weeks of gestation and averaged over the two time points. In addition, seven OP insecticide metabolites, including TCPy, MDA, 4-nitrophenol (PNP; a metabolite of methyl parathion, parathion, and other chemicals), 2-diethylamino-4-hydroxy-6-methylpyrimidine (a metabolite of pirimiphos methyl), 2-isopropyl-4-methyl-6 -hydroxypyrimidine (a metabolite of diazinon), 3-chloro-4-methyl-7-hydroxycoumarin (a metabolite of coumaphos and coumaphos methyl), and 5-chloro-1-isopropyl-3-hydroxytriazole (a metabolite of isazophos and isazophos methyl), were measured in the same maternal prenatal urine specimens. The last four were detected in fewer than 11% of subjects and therefore were not studied further. Median concentrations (range) in maternal prenatal urine were 136 nmol/L (10–6854) for DAPs, 101 nmol/L (5–6587) for DMPs, 22 nmol/L (2–680) for DEPs, 3.3 μg/L (0.2–56.1) for TCPy, 0.2 μg/L (0.2–28.9) for MDA, and 0.5 (0.1–34.7) for PNP (Eskenazi et al. 2004). AChE and BChE activities were measured in maternal blood taken at ~26–27 weeks of gestation and before delivery, and in umbilical cord blood taken at delivery. In addition, $PON1_{192}$ and $PON1_{-108}$ polymorphisms were genotyped in maternal and cord blood specimens, and PON1 arylesterase activity against phenyl acetate (as a measure of PON1 quantity) and paraoxonase activity against paraoxon (as a measure of PON1 activity) were measured in maternal post-delivery and cord blood specimens.

After multivariate adjustment in 485 mother–newborn pairs, a 10-fold increase (i.e., 1 $\log_{10}$-nmol/L increase) in maternal prenatal urinary DAP concentration was positively, but not significantly, associated with birth length (beta = 0.52 cm, 95% CI = −0.01, 1.05) and was significantly positively associated with head circumference (beta = 0.32 cm, 95% CI = 0.03, 0.62) (Table 2) (Eskenazi et al. 2004). After urinary DAP levels were controlled for creatinine, the association with birth length was no longer evident; however, the result for head circumference did not change. Maternal prenatal urinary DAP concentrations were not significantly associated with length of gestation, preterm delivery (birth at <37 weeks of gestation), birth weight, low birth weight (<2500 g), ponderal index, or small size for gestational age at birth (<10th percentile for birth weight at gestational age), nor were maternal prenatal urinary DMP or DEP concentrations significantly associated with these outcomes, other than an inverse association between prenatal DMPs and length of gestation (beta = −0.41 weeks, 95% CI = −0.75, −0.07). In analyses by timing of prenatal DMP measurement, the latter association appeared to be stronger after 22 weeks of gestation. Maternal prenatal urinary TCPy and MDA levels were not significantly associated with any outcome evaluated. When maternal prenatal urinary PNP levels were categorized as undetectable, detectable, and below the median, or detectable and at or above the median, newborns in the middle category of exposure, but not the highest category, had a shorter length of gestation and longer birth length than those in the lowest category. However, the authors cautioned that "PNP may derive from compounds other than parathion" (Eskenazi et al. 2004). AChE activities in cord blood (beta = 0.34 weeks, 95% CI = 0.13, 0.55) were significantly associated with longer gestation, and lower levels were associated with a significantly higher odds of preterm birth and low birth weight. However, activities in maternal prenatal and delivery blood were not significantly associated with length of gestation, birth weight, or the other outcomes assessed. Activities of BChE in maternal prenatal plasma, maternal plasma at delivery, and cord plasma at delivery also were not significantly associated with any of the outcomes examined. The authors did not have baseline AChE data; thus, AChE and BChE inhibition was not measured. Furthermore, AChE and BChE levels may vary significantly across time due to changes in OP insecticide exposure and/or natural variability. In stratified analyses by *PON1* genotype or PON1 activity, maternal prenatal urinary DEP levels were associated with shorter gestational age only among infants with the $PON1_{-108}$ TT (low-activity) genotype ($P_{interaction} = 0.09$; Table 2) (Harley et al. 2011). Maternal prenatal urinary DAP and DMP levels were (non-significantly) associated with higher birth weight only among those with the $PON1_{-108}$ CT genotype ($P_{interaction} = 0.06$ and 0.05, respectively), whereas associations with birth weight were statistically non-significant among those with the TT genotype. Positive associations with head circumference were detected only among those with $PON1_{-108}$ CT or $PON1_{192}$ RR (low-activity) genotype. Cord blood PON1 arylesterase and paraoxonase activity levels were not significant modifiers of the associations between maternal prenatal DAP, DMP, or DEP concentrations and birth outcomes, although a significant positive association between prenatal DEPs and birth weight was detected only among those with high cord blood levels of PON1.

The CHAMACOS study has several methodological strengths, including its relatively large size, evaluation of numerous potential confounders, and collection of several individual-level OP metabolites around the beginning of the second and third trimesters of pregnancy.

Limitations of CHAMACOS include use of on only two averaged biospecimens to characterize exposure and the other concerns identified above for the CCCEH and CECS studies, as well as the inherent shortcomings of DAP metabolites as biomarkers of OP insecticide exposure. Additionally, the 53.2% participation rate (with a follow-up rate of approximately 88%) raises concerns about selection bias, although the direction and magnitude of such bias cannot be quantified. The main analyses in the whole CHAMACOS cohort indicate that maternal prenatal levels of DAPs, DMPs, DEPs, TCPy, and MDA, and activities of AChE or BChE, were not associated with most birth outcomes examined. The only exceptions were the positive association between DAPs and head circumference and the inverse association between DMPs and length of gestation, especially when DMP concentrations were measured after the midpoint of pregnancy. The associations of maternal prenatal PNP levels with shorter gestation and greater body length were not consistent with a monotonic exposure–response trend, and the findings for AChE activities in cord blood were not consistent with the findings for activities in maternal prenatal and perinatal blood. The observed associations were not modified by PON1 quantity or activity, but some evidence of modification by *PON1* genotype was found, albeit with somewhat contradictory patterns (e.g., positive associations with birth weight in $PON1_{-108}$ CT carriers, but positive associations with

head circumference in $PON1_{192}$ RR carriers). The authors not only interpreted the inverse association between DMPs and length of gestation as being consistent with a stimulatory effect of OP insecticides on uterine contraction, but also noted that the 6.4% rate of preterm delivery in this population was lower than the U.S. average (Eskenazi et al. 2004). Information is lacking on effects of OP insecticides on uterine smooth muscle. However, in mouse uterus, regulation of acetylcholine levels is dominated by BChE and activity changes in excess of 50%, which can occur during the estrus cycle, appear to be required to cause substantial changes in uterine contractile activity (Medina et al. 1993). In light of the scattered positive associations with some but not all indicators of fetal growth and the internally inconsistent associations with length of gestation, the overall results do not demonstrate an adverse effect of prenatal exposure to OP insecticides on birth outcomes.

New Jersey birth cohort

In a convenience sample of 150 New Jersey women with a non-anomalous singleton pregnancy scheduled for an elective cesarean birth at ≥37 weeks of gestation in 2003–2004, chlorpyrifos and other pesticides were measured in preoperative maternal serum and umbilical cord serum (Table 1) (Barr et al. 2010). The mean chlorpyrifos concentration was 0.09 ng/g (SD = 0.87, range = 0.0007–10.09) in maternal serum and 0.55 ng/g (SD = 0.73, range = 0.0007–1.84) in cord serum. After multivariate adjustment, mean birth weight, head circumference, abdominal circumference, and birth length did not differ significantly between newborns with maternal prenatal or cord serum chlorpyrifos concentrations ≥75th versus <75th percentile (0.0007 ng/g), nor did they differ significantly when the cutoff was set at the 90th percentile (Table 2).

Strengths and limitations of the New Jersey birth cohort study are comparable to those described above for other prospective birth cohort studies. Chlorpyrifos levels measured in maternal blood just before cesarean section may not be accurate indicators of earlier prenatal levels, which are probably more etiologically relevant to fetal growth. Because subjects were recruited by convenience sampling, the participation rate was not stated, and selection bias is a possibility if participation was related to factors associated with both chlorpyrifos exposure and birth outcomes (e.g., socioeconomic status, diet, and place of residence). It is unclear whether results for newborns delivered by elective cesarean section would be expected to differ from those for newborns delivered vaginally or by unplanned cesarean section. Finally, the scope of the study with regard to OP insecticides was limited by the measurement of only chlorpyrifos. As a whole, the results of this study do not demonstrate an association between detectable chlorpyrifos in maternal perinatal serum and birth outcomes.

Shanghai birth cohort

Among 187 healthy women in Shanghai with an uncomplicated singleton pregnancy in 2006–2007, five DAP metabolites were measured in maternal spot urine specimens collected at the onset of labor (Table 1) (Wang et al. 2012). Geometric mean concentrations were 17.19 μg/L (range = < LOD–269.15) for DMP, 8.01 μg/L (range = <LOD–109.65) for DMTP, 6.03 μg/L (range = <LOD–109.65) for DEP, 6.31 μg/L (range = < LOD–131.83) for DETP, and undetectable (range = < LOD–5.1; 5.34% detectable) for DEDTP. In multivariate adjusted models including all 187 newborns, no significant association was detected between any maternal prenatal urinary DAP metabolite or all DAPs combined and length of gestation, birth weight, or body length (Table 2). Among the 91 girls, log-transformed DEP concentration was significantly inversely associated with length of gestation (beta = −1.79 weeks, 95% CI = −2.82, −0.76; log scale not specified), but no such association was observed among boys (beta = 0.17 weeks, $P = 0.164$). No other significant associations were reported among girls or boys only.

The strengths and limitations of the Shanghai birth cohort study have been described in the context of other prospective birth cohort studies. The high participation rate (stated as 97% among eligible subjects) is a strength, although the derivation of this rate (i.e., the definition of the eligible source population) is not clear.

As stated earlier, DAP metabolite levels measured at the time of labor may not reflect earlier exposure levels, which may be more relevant to fetal development. In addition, the limitations of DAP metabolites for OP exposure assessment were discussed previously. The authors noted that DAP metabolite levels observed in this study were substantially higher than those reported among pregnant or postpartum women in the United States (Bradman et al. 2005) and the Netherlands (Ye et al. 2008), yet only one statistically significant association was detected. The authors did not hypothesize why levels of DEP, but not other DAP metabolites, might plausibly be related to shorter length of gestation, but not other birth outcomes, only among girls. With at least 54 associations tested, chance appears to be a more reasonable explanation for this single statistically significant result.

Health Outcomes and Measures of the Environment Study

In the Health Outcomes and Measures of the Environment (HOME) Study, based in Cincinnati, Ohio, 389 healthy pregnant women living in a home built before 1978 were enrolled at ≤19 weeks of gestation and followed through delivery of a live-born singleton infant in 2003–2006 (Table 1) (Rauch et al. 2012). Six DAP metabolites were measured in maternal spot urine samples collected from 344 participants at approximately 16 and 26 weeks of gestation (averaged for analysis) and within 24 h of delivery. Median concentrations were 81.3 nmol/L (IQR = 41.7–220.0) for DAPs, 56.9 nmol/L (IQR = 26–185) for DMPs, and 17.7 nmol/L (IQR = 8–37) for DEPs. In addition, umbilical cord blood was genotyped for the $PON1_{192}$ and $PON1_{-108}$ polymorphisms.

Statistically significant inverse associations were detected in multivariate adjusted models between $\log_{10}$-transformed, creatinine-standardized maternal prenatal urinary DAP or DMP concentrations and length of gestation (beta for DAPs = −0.5 weeks, 95% CI = −0.8, −0.1; beta for DMPs = −0.4 weeks, 95% CI = −0.7, 0.0) and birth weight (beta for DAPs = −151 g, 95% CI = −287, −16; beta for DMPs = −124 g, 95% CI = −245, −2) (Table 2) (Rauch et al. 2012). However, the associations with birth weight were substantially attenuated and not statistically significant after adjustment for gestational age. Maternal prenatal

urinary DEP concentrations were not significantly associated with either outcome. After stratification by race, the inverse associations of prenatal DAPs and DMPs with length of gestation were detected only for white mothers ($P_{interaction} = 0.10$ and 0.09, respectively), whereas the inverse associations of prenatal DAPs and DMPs with birth weight were detected only for black mothers and only after additionally adjusting for gestational age ($P_{interaction} = 0.02$ and 0.02, respectively). In models stratified by genotype, maternal prenatal urinary DAP concentrations were inversely associated with length of gestation only among newborns with the $PON1_{192}$ QR or QQ (not the low-activity RR) genotype ($P_{interaction} = 0.04$ and 0.09, respectively) or the $PON1_{-108}$ CT (not the high-activity CC or low-activity TT) genotype ($P_{interaction} = 0.04$). The inverse association between maternal prenatal urinary DAP concentration and birth weight was observed only among newborns with the $PON1_{192}$ QR genotype ($P_{interaction} = 0.02$) or the $PON1_{-108}$ CT genotype ($P_{interaction} = 0.15$). Models stratified by both race and genotype were not only based on small numbers and were therefore statistically unstable, with wide CIs, but also suggested stronger associations among heterozygotes. Results were modestly attenuated when restricted to full-term births, based on non-creatinine-standardized DAP concentrations, or based on DAP concentrations from either 16 or 26 weeks only.

Besides the previously noted strengths and limitations of prospective birth cohort studies, this study benefits from the measurement of maternal urinary DAP metabolite levels at two time points near the beginning and end of the second trimester. The 37.1% participation rate among 1263 eligible women, combined with the 86% follow-up rate through delivery (including twins and stillbirths) and the 88% biospecimen availability rate, raises the possibility of selection bias, with an unknown direction and magnitude of influence. Only DAP metabolites, not specific OP insecticides, were measured, and results were reported only for two birth outcomes. Several anomalous results were reported, including the detection of some associations only among white mothers and others only among black mothers; the detection of significant inverse associations between maternal prenatal urinary DAP and DMP levels and birth weight only in the absence of adjustment for gestational age, but the detection of these associations among black mothers only with adjustment for gestational age; and the stronger observed associations among $PON1_{192}$ and $PON1_{-108}$ heterozygotes than among low-activity homozygotes. Overall, the results suggest possible inverse associations between maternal prenatal urinary DAP and DMP (but not DEP) concentrations and length of gestational age and birth weight, perhaps restricted to specific racial groups or those with intermediate (but not low or high) PON1 activity. As in other studies with internally inconsistent associations, these findings may be explained at least in part by chance or bias, especially given the numerous hypotheses tested, and do not offer convincing support for an adverse effect of OP insecticides on birth outcomes.

Zhejiang birth cohort

In rural Zhejiang Province, 116 consecutive healthy women with a healthy, uncomplicated, singleton pregnancy at 36 weeks of gestation, eight OP insecticides (and other pesticides) were measured in umbilical cord serum at delivery, including chlorpyrifos, diazinon, fonofos, malathion, parathion, methylparathion, profenofos, and terbufos (Table 1) (Wickerham et al. 2012). The proportion of serum samples with detectable levels (LOD = 0.05 ng/mL except for malathion and profenofos, where LOD = 0.50 ng/mL) were 23.3% for chlorpyrifos (90th percentile = 0.17 ng/mL), 14.7% for diazinon (90th percentile = 0.27 ng/mL), 16.4% for fonofos (90th percentile = 0.30 ng/mL), 25.9% for malathion (90th percentile = 3.13 ng/mL), 2.6% for parathion (90th percentile $<$ 0.05 ng/mL), 25.0% for profenofos (90th percentile = 0.68 ng/mL), and 31.0% for terbufos (90th percentile = 0.27 ng/mL). After multivariate adjustment, no significant associations with birth weight were detected for any of these pesticides, whether analyzed as detectable versus non-detectable, three-level ordinal variables, or the total number detected (Table 2).

In general, some of the strengths and limitations of the Zhejiang birth cohort study are similar to those of other birth cohort studies. Additional strengths include the nearly 100% participation rate among eligible subjects (although the basis for calculating this rate is not clear) and the measurement of specific OP insecticides rather than non-specific DAP metabolites, countered by limitations, including the cross-sectional measurement of pesticide levels in umbilical cord serum collected at delivery, the analysis of pesticide exposure as simplified categorical variables, and the evaluation of only a single birth outcome (weight). In summary, the results of this study suggest no association between concurrent exposure to any of eight different OP insecticides and birth outcomes.

Bradford Hill evaluation of weight of evidence

Strength. The strength of observed associations between OP metabolites and birth outcomes cannot be compared easily across studies, given differences in the unit of exposure measurement, the logarithmic base used for transformation (if any), and the format in which results were presented (e.g., as regression betas or adjusted means). The distinction between a weak and a strong association, especially with a continuous outcome such as birth weight or length of gestation, is also subjective and hard to define. Nevertheless, most reported associations involved birth weight differences of $<$ 100 g, birth length and head circumference differences of $<$0.5 cm, and gestational length differences of $<$5 days ($<$0.7 weeks), in association with exposures classified on various scales (e.g., detectable, natural log, or $\log_{10}$). In general, weak associations are more likely than strong associations to be explained by confounding, bias, or chance. Furthermore, the majority of reported results were not statistically significantly different from the null value.

Consistency. An examination of the consistency of associations with specific birth outcomes reveals mostly null findings, with no uniformity of positive or inverse associations across (as well as within) studies. In particular, associations with birth weight were inconsistently reported as inverse (Perera et al. 2003, Rauch et al. 2012, Whyatt et al. 2005, Whyatt et al. 2004), positive (Eskenazi et al. 2004, Harley et al. 2011), or in most cases, null (Barr et al. 2010, Berkowitz et al. 2004, Eskenazi et al. 2004, Harley et al. 2011, Perera et al. 2003, Wang et al. 2012, Whyatt et al. 2005, Whyatt et al. 2004, Wolff

et al. 2007) (Wickerham et al. 2012). Associations with birth length were also heterogeneous, including inverse (Perera et al. 2003, Whyatt et al. 2005, Whyatt et al. 2004), positive (Eskenazi et al. 2004), and mostly null findings (Barr et al. 2010, Berkowitz et al. 2004, Eskenazi et al. 2004, Perera et al. 2003, Wang et al. 2012, Whyatt et al. 2005, Whyatt et al. 2004, Wolff et al. 2007). Likewise, head circumference was variously inversely associated (Berkowitz et al. 2004, Wolff et al. 2007), positively associated (Eskenazi et al. 2004, Harley et al. 2011), and most often not significantly associated (Barr et al. 2010, Berkowitz et al. 2004, Eskenazi et al. 2004, Harley et al. 2011, Perera et al. 2003, Whyatt et al. 2005, Whyatt et al. 2004, Wolff et al. 2007) with OP metabolite levels. The reported associations of individual OP metabolites with ponderal index were statistically null in both studies of this outcome (Eskenazi et al. 2004, Wolff et al. 2007).

Although inverse associations between a few selected OP metabolites and length of gestation were reported in multiple studies (Eskenazi et al. 2004, Harley et al. 2011, Rauch et al. 2012, Wang et al. 2012), these associations were inconsistent across participant subgroups by race, sex, and genotype, with no cogent biological explanation for the observed heterogeneity. For example, a significant inverse association with maternal prenatal urinary DAPs was detected among white mothers but not black mothers in the HOME Study (Rauch et al. 2012), and an inverse association with maternal prenatal urinary DEPs was detected among infant girls but not boys in the Shanghai birth cohort study (Wang et al. 2012). Moreover, most OP metabolites measured in these and other studies (Berkowitz et al. 2004, Wolff et al. 2007) were not significantly associated with length of gestation. Thus, the most consistent findings were statistically null, and the lack of consistency of significant associations between OP metabolites and specific birth outcomes does not support a causal interpretation of the few statistically significant associations observed.

Temporality. As discussed above, an assessment of the temporal relationship of measured OP and DAP metabolite levels in prenatal or perinatal maternal biospecimens in relation to fetal growth and other birth outcomes is limited, for several reasons. First, exposures measured soon before birth are unlikely to have a major influence on fetal growth over the course of 40 weeks of gestation. Second, because these metabolites have a short biological half-life and vary considerably within individuals, one or two samples are unlikely to reflect past or long-term average exposure for a given person. Third, it is unknown whether the time points selected for blood, urine, or personal air collection in various studies are etiologically relevant or whether exposures earlier or later in gestation have a greater influence on fetal growth or length of gestation. Consequently, the measurement of metabolite levels prior to birth does not necessarily strengthen the evidence in favor of a causal interpretation of observed associations, especially if measurements were taken only hours or minutes before birth, but even if they were made months in advance.

Biological gradient. Few studies explicitly evaluated the shape of the biological gradient between OP metabolites and birth outcomes; instead, using linear regression models, nearly all investigators assumed a log-linear exposure–outcome relationship, without testing the appropriateness of this model. Only two studies examined exposure–response gradients by categorizing exposures into at least three ordinal groups (Eskenazi et al. 2004, Whyatt et al. 2004). (This observation also highlights the problem of inconsistency of analytic approaches among studies.) In the CCCEH study, where cord plasma concentrations of chlorpyrifos (and chlorpyrifos plus diazinon, but not diazinon alone) were inversely associated with birth weight and birth length, the strength of the inverse associations increased across tertiles of detectable levels compared with non-detectable levels, a pattern consistent with a monotonic exposure–response gradient (Whyatt et al. 2004). In the CHAMACOS study, maternal prenatal urinary levels of MDA, TCPy, and PNP, which were categorized as undetectable, detectable below the median, or detectable at or above the median, did not show evidence of a monotonic association with length of gestation, birth weight, birth length, or ponderal index (Eskenazi et al. 2004). The middle category of PNP appeared to be inversely associated with length of gestation and ponderal index and positively associated with body length; however, the results for the highest category were not significantly different from the null value. Some evidence of a positive exposure–response trend was observed between PNP and head circumference. Although significant linear regression coefficients may be consistent with a monotonic biological gradient, the dearth of information on the shape of exposure–response relationships between OP metabolites and birth outcomes prevents a thorough evaluation of such gradients.

The commonly applied mechanism for OP toxicity is AChE inhibition. As discussed earlier, the levels in the epidemiologic studies are orders of magnitude below what would result in clinically meaningful AChE inhibition. There are a few other postulated mechanisms for non-cholinergic OP toxicity, but effects at the levels observed in the epidemiologic studies have not been established for these mechanisms either.

Plausibility. The biological plausibility of the associations is not established. While OP insecticides are known to cause neurotoxicity in mature subjects at doses higher than reported in the epidemiologic studies, the mechanism of OP-induced neurodevelopmental toxicity has yet to be established.

Coherence. In the evaluation of the coherence of evidence, another important consideration is whether observed interactions with PON1 activity levels or genotypes are consistent with the hypothesis of increased susceptibility to potential adverse health effects of OP insecticides in those with lower PON1 activity. In the three studies that evaluated these interactions—the Mount Sinai CECS (Berkowitz et al. 2004, Wolff et al. 2007), CHAMACOS (Harley et al. 2011), and the HOME Study (Rauch et al. 2012)—results were variable. One study reported the expected stronger inverse associations, albeit only between selected metabolites and birth outcomes, in those with homozygous low-activity PON1 genotypes or low measured PON1 activity (Wolff et al. 2007). Another study found mostly no apparent heterogeneity by PON1 genotype, level, or activity, but some evidence of stronger positive associations, again between only selected metabolites and birth outcomes, in those with heterozygous or homozygous high-activity PON1 genotypes or higher PON1 levels, and an inverse association

between DEPs and gestational age among those homozygous for the low-activity $PON1_{-108}$ genotype (Harley et al. 2011). In another study, stronger inverse associations between DAPs and birth outcomes were observed among PON1 heterozygotes than among low- or high-activity homozygotes (Rauch et al. 2012). Finally, one study found no evident heterogeneity in associations by PON1 activity (Berkowitz et al. 2004). As a whole, these mixed results are not coherent with a protective effect of high PON1 detoxifying activity against adverse effects of OP insecticides on fetal growth and other birth outcomes.

Specificity, experiment, and analogy. The other Bradford Hill guidelines—specificity, experiment, and analogy—are less informative for the evaluation of causality. Especially in light of the non-specificity of DAP metabolites, the many influences on birth outcomes, and the numerous associations tested, no specific relationship has emerged between any particular OP insecticide and any particular birth outcome. Relevant quasi-experimental evidence in humans, such as a study of birth outcomes in women who adhere to an organic diet, is unavailable. Drawing analogies with other prenatal exposures that cause adverse birth outcomes (e.g., ethanol, methylmercury, certain prescription medications, and dietary factors) is not warranted, just like existence of numerous exposures shown to be safe cannot be used to refute a causal association between OP insecticides and adverse birth outcomes. On balance, such analogies do not sway the evaluation of causality.

Neurodevelopmental outcomes

Twenty studies in ten study populations have examined associations between OP or OP metabolites and neurodevelopmental outcomes (Bouchard et al. 2010, Bouchard et al. 2011, Engel et al. 2007, Engel et al. 2011, Eskenazi et al. 2010, Eskenazi et al. 2007, Fortenberry et al. 2014, Guodong et al. 2012, Horton et al. 2012, Lizardi et al. 2008, Lovasi et al. 2011, Marks et al. 2010, Oulhote and Bouchard 2013, Quiros-Alcala et al. 2011, Rauh et al. 2011, Rauh et al. 2006, Rauh et al. 2012, Yolton et al. 2013, Young et al. 2005, Zhang et al. 2014). Most studies were conducted in birth cohorts enrolled prior to delivery—including four cohorts described earlier in the section on birth outcomes—whereas other studies were cross-sectional in design. Measures of neurodevelopment varied among studies, with several using standard clinical scales or questionnaires, and others using measurement tools that were not used by any other studies reviewed, although all studies reported some degree of validation of the assessment tools used. Table 3 summarizes the analyses in the studies evaluating neurodevelopmental outcomes.

Columbia Center for Children's Environment and Health

In the CCCEH birth cohort study, which was described earlier with respect to birth outcomes, childhood neurodevelopmental outcomes were measured using the Bayley Scales of Infant Development, 2nd Edition, including the Mental Development Index and the Psychomotor Development Index, to assess cognitive and psychomotor development at ages 12, 24, and 36 months; the mother-reported Child Behavior Checklist for ages 1.5–5 years, including syndrome scale scores, internalizing and externalizing scores, and scales oriented to the *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition* (DSM-IV), to assess recent behavioral problems at age 36 months; the Child Behavior Checklist for ages 6–18 years to assess recent behavioral problems at age 7 years; the Wechsler Intelligence Scale for Children, 4th Edition, including the Verbal Comprehension Index, the Perceptual Reasoning Index, the Working Memory Index, and the Processing Speed Index, which were combined to determine the Full-Scale Intelligence Quotient (IQ), at age 7 years; and magnetic resonance imaging for brain morphology at ages 5.9–11.2 years (Table 1) (Horton et al. 2012, Lovasi et al. 2011, Rauh et al. 2011, Rauh et al. 2006, Rauh et al. 2012).

After multivariate adjustment, significant or borderline significant inverse associations were observed between the highest detectable tertile (>6.17 pg/g) versus lower levels of cord plasma chlorpyrifos and the Bayley Mental Development Index (beta = −3.327, SE = 1.76, $P = 0.06$) and the Psychomotor Development Index (beta = −6.46, SE = 2.18, $P = 0.003$) at age 36 months (Table 2) (Rauh et al. 2006). The inverse association with the Mental Development Index at 36 months was observed only among African American children (beta = −6.34), and not among Dominican children (beta = −1.70), whereas the inverse association with the Psychomotor Development Index at 36 months was observed in both groups (beta = −7.15 and −5.18, respectively). No interactions were detected with other covariates tested. When the Bayley indices were dichotomized at 85 points (one SD below the mean) to indicate developmental delay, cord plasma chlorpyrifos levels in the highest detectable tertile were associated with a significantly increased odds of mental delay (odds ratio [OR] = 2.37, 95% CI = 1.08, 5.19) and psychomotor delay (OR = 4.52, 95% CI = 1.61, 12.70) at 36 months. However, cord plasma chlorpyrifos was not significantly associated with the Mental Development Index at 12 or 24 months (beta at 12 months = −0.344, SE = 1.66; beta at 24 months = −1.480, SE = 2.03) or with the Psychomotor Development Index at either time point (beta at 12 months = −3.30, SE = 2.11; beta at 24 months = 1.17, SE = 1.98), nor were significant associations detected with mental or psychomotor delay at those ages. At 36 months, significant associations were detected between elevated chlorpyrifos levels and Child Behavior Checklist measures of attention problems (OR = 11.26, 95% CI = 1.79, 70.99), attention deficit/hyperactivity disorder (ADHD; OR = 6.50, 95% CI = 1.09, 38.69), and pervasive developmental disorder (OR = 5.39, 95% CI = 1.21, 24.11), but not externalizing behavior problems (unadjusted $P = 0.426$) or internalizing behavior problems (unadjusted $P = 0.444$). A subsequent analysis of neighborhood characteristics based on U.S. census data for poverty, education, race, language, and housing showed no substantial confounding (<10% change in beta) or modification ($P \geq 0.20$) of the associations between cord plasma chlorpyrifos and the Bayley Mental Development and Psychomotor Indices at 36 months (Table 2) (Lovasi et al. 2011).

Based on results of the Wechsler Intelligence Scale testing at age 7 years, with outcomes analyzed on the natural log scale, no significant adjusted associations were detected between cord plasma chlorpyrifos levels and Wechsler Full-Scale IQ, Verbal Comprehension, Perceptual Reasoning, or Processing Speed, nor were any significant interactions with

Table 3. Results of epidemiologic studies of organophosphorus insecticide biomarkers and neurodevelopmental outcomes.

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Rauh et al. (2006)	Bayley Mental Development Index at 12, 24, or 36 months	Cord plasma chlorpyrifos >6.17 vs. ≤6.17 pg/g	254 total 229 at 12 months 225 at 24 months 228 at 36 months	Beta ± SE at 12 months = –0.344 ± 1.66, $P = 0.836$ Beta ± SE at 24 months = –1.480 ± 2.03, $P = 0.466$ Beta ± SE at 36 months = –3.327 ± 1.76, $P = 0.060$ Beta at 36 months, African Americans = –6.34, $P < 0.05$ Beta at 36 months, Dominicans = –1.70, $P \geq 0.05$ "Interaction terms for the interaction of chlorpyrifos exposure with the other exposure and sociodemographic variables were tested in the full model, and none was significant." Generalized linear models showed no significant within-subject association with chlorpyrifos over age groups ($P = 0.23$)	Prenatal environmental tobacco smoke, race/ethnicity, infant gender, maternal intelligence quotient by Test of Nonverbal Intelligence (Second Edition), maternal education, and Home Observation for Measurement of the Environment score	Cutoff for cord plasma chlorpyrifos was set at the highest tertile of detectable levels because "the only group for which mean 36-month [Bayley Scales of Infant Development II] scores were significantly lower was the group with the highest exposure level (> 6.17 pg/g)." Bayley scores are standardized to a mean ± SD of 100 ± 15, with scores ≤ 85 indicating developmental delay (minimum score = 50, maximum score = 150) "When administered at 3 years of age, the [Bayley Scales of Infant Development II] demonstrates only moderate predictive power for subsequent intelligence and school performance but is clinically useful for children performing in the subnormal range."
Rauh et al. (2006)	Bayley Psychomotor Development Index at 12, 24, or 36 months	"	"	Beta ± SE at 12 months = –3.30 ± 2.11, $P = 0.12$ Beta ± SE at 24 months = 1.17 ± 1.98, $P = 0.56$ Beta ± SE at 36 months = –6.46 ± 2.18, $P = 0.003$ Beta at 36 months, African Americans = –7.15, $P < 0.05$ Beta at 36 months, Dominicans = –5.18, $P < 0.05$ "All interaction terms for the interaction of chlorpyrifos exposure with the other exposure and sociodemographic variables were tested in the full model, and none was significant." Generalized linear models showed a significant within-subject association with chlorpyrifos over age groups ($P = 0.01$), with a difference emerging between 24 and 36 months ($P = 0.003$)	"	—

Rauh et al. (2006)	Bayley mild/ significant mental delay at 12, 24, or 36 months	"	"	Odds ratio at 12 months = 1.22 (0.48, 3.06) Odds ratio at 24 months = 1.75 (0.86, 3.60) Odds ratio at 36 months = 2.37 (1.08, 5.19)	"	–
Rauh et al. (2006)	Bayley mild/ significant psychomotor delay at 12, 24, or 36 months	"	"	Odds ratio at 12 months = 1.88 (0.78, 4.53) Odds ratio at 24 months = 1.01 (0.37, 2.76) Odds ratio at 36 months = 4.52 (1.61, 12.70)	"	–
Rauh et al. (2006)	Child Behavior Checklist attention problems at 36 months	"	228	Odds ratio = 11.26 (1.79,70.99)	"	Child Behavior Checklist collects information on behaviors occurring in the past 2 months, with the cutoff for borderline or clinical problems set at 98th percentile
Rauh et al. (2006)	Child Behavior Checklist ADHD problems at 36 months	"	"	Odds ratio = 6.50 (1.09, 38.69)	"	–
Rauh et al. (2006)	Child Behavior Checklist pervasive developmental disorder problems at 36 months	"	"	Odds ratio = 5.39 (1.21, 24.11)	"	–
Rauh et al. (2006)	Child Behavior Checklist externalizing behavior problems at 36 months	"	"	% >6.17 = 10.6% % ≤6.17 = 8.6% $P = 0.426$	None	–
Rauh et al. (2006)	Child Behavior Checklist internalizing behavior problems at 36 months	"	"	% >6.17 = 14.9% % ≤6.17 = 13.0% $P = 0.444$	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Lovasi et al. (2011)	Bayley Mental Development Index at 36 months	Cord plasma chlorpyrifos > 6.17 vs. ≤ 6.17 pg/g	266	Model 1 beta = –3.2 (–5.1, –1.3) Model 2 beta = –3.4 (–5.2, –1.5) Model 3 beta = –3.2 (–5.0, –1.5) Model 4 beta = –3.1 (–4.8, –1.3) Model 5 beta = –3.0 (–4.8, –1.2) Model 6 beta = –3.2 (–5.1, –1.3)	Infant gender, gestational age, Dominican ethnicity, maternal education, maternal intelligence quotient, prenatal environmental tobacco smoke exposure, and index of building disrepair plus: Model 1: none additional Model 2: neighborhood % poverty and % high school graduates Model 3: neighborhood % African American Model 4: neighborhood % linguistic isolation Model 5: neighborhood % crowded household Model 6: neighborhood inadequate plumbing and % vacant housing	Residential neighborhoods characterized by mothers' self-report and U.S. Census data within geocoded network buffers Neighborhood poverty did not significantly modify the association of chlorpyrifos exposure with Bayley Mental Development Index ($P = 0.2$)
Lovasi et al. (2011)	Bayley Psychomotor Development Index at 36 months	"	"	Model 1 beta = –6.9 (–11.1, –2.7) Model 2 beta = –7.0 (–11.0, –2.9) Model 3 beta = –7.3 (–11.5, –3.0) Model 4 beta = –7.2 (–11.3, –3.0) Model 5 beta = –6.9 (–11.1, –2.8) Model 6 beta = –7.1 (–11.4, –2.7)	"	Neighborhood poverty did not significantly modify the association of chlorpyrifos exposure with Bayley Psychomotor Development Index ($P = 0.4$)
Rauh et al. (2011)	Wechsler full-scale intelligence quotient at 7 years, natural log scale	Cord plasma chlorpyrifos (pg/g)	265	Parsimonious model beta = –0.003 (–0.006, 0.001) Fully adjusted beta = –0.003 (–0.006, 0.000) Change per SD (4.61 pg/g) increase in exposure = –1.4% "No significant interactions" between chlorpyrifos and any covariates	Parsimonious model (least absolute shrinkage and selection operator): maternal education, maternal intelligence quotient, and Home Observation for Measurement of the Environment score Fully adjusted model: child sex, race/ethnicity, maternal intelligence quotient, maternal education, household income, child age at testing, prenatal environmental tobacco smoke exposure, and prenatal polycyclic aromatic hydrocarbons exposure	Full-scale intelligence quotient is the sum of four composite indices; mean ± SD = 100 ± 15 $P = 0.08$ for smoothed cubic spline model vs. linear model for chlorpyrifos (unadjusted)
Rauh et al. (2011)	Wechsler verbal comprehension at 7 years, natural log scale	"	"	Parsimonious model beta = none; chlorpyrifos dropped from model Fully adjusted beta = –0.002 (–0.005, 0.001) "No significant interactions" between chlorpyrifos and any covariates	"	Verbal comprehension index measures verbal concept formation and predicts school readiness; mean ± SD = 100 ± 15 $P = 0.07$ for smoothed cubic spline model vs. linear model for chlorpyrifos (unadjusted)

Rauh et al. (2011)	Wechsler perceptual reasoning at 7 years, natural log scale	"	"	Parsimonious model beta = none; chlorpyrifos dropped from model Fully adjusted beta = –0.002 (–0.006, 0.002) "No significant interactions" between chlorpyrifos and any covariates	"	Perceptual reasoning index measures nonverbal and fluid reasoning; mean ± SD = 100 ± 15 $P = 0.08$ for smoothed cubic spline model vs. linear model for chlorpyrifos (unadjusted)
Rauh et al. (2011)	Wechsler processing speed at 7 years, natural log scale	"	"	Parsimonious model beta = none; chlorpyrifos dropped from model Fully adjusted beta = 0.001 (–0.004, 0.005) "No significant interactions" between chlorpyrifos and any covariates	"	Processing speed index assesses ability to focus attention and quickly scan, discriminate, and sequentially order visual information; mean ± SD = 100 ± 15 $P = 0.59$ for smoothed cubic spline model vs. linear model for chlorpyrifos (unadjusted)
Rauh et al. (2011)	Wechsler working memory at 7 years, natural log scale	"	"	Parsimonious model beta = –0.006 (–0.009, –0.002) Fully adjusted beta = –0.006 (–0.010, –0.002) Change per SD (4.61 pg/g) increase in exposure = –2.8% "No significant interactions" between chlorpyrifos and any covariates	"	Working memory index assesses ability to memorize new information, hold it in short-term memory, concentrate, and manipulate information; mean ± SD = 100 ± 15 $P = 0.40$ for smoothed cubic spline model vs. linear model for chlorpyrifos (unadjusted)
Horton et al. (2012)	Wechsler working memory at 7 years	Cord plasma chlorpyrifos (pg/g, natural log scale)	335	Model 0 beta, males = − 2.382 (− 3.88, − 0.88) Model 0 beta, females = − 0.524 (− 1.90, 0.85) Model 1 beta = –1.451 (− 2.265, − 0.438) Model 2 beta = − 1.355 (− 2.368, − 0.341) Model 3 beta = − 1.478 (− 2.496, − 0.459)	Model 0: None Family income, maternal education, and child sex plus: Model 1: total Home Observation for Measurement of the Environment score Model 2: parental nurturance composite scale of the Home Observation for Measurement of the Environment score Model 3: environmental stimulation composite scale of the Home Observation for Measurement of the Environment score	Home Observation for Measurement of the Environment score based on evaluation of child home environment at age 3 years Parental nurturance: sum of z-scores of responsivity, modeling, and acceptance subscales, which measure such maternal behaviors as attentiveness, displays of physical affection, encouragement of delayed gratification, limit setting, and the ability of the mother to control her negative reactions Environmental stimulation: sum of z-scores of learning materials, language stimulation, academic stimulation, and variety subscales, which measure the availability of intellectually stimulating materials in the home and the mother's encouragement of learning

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Horton et al. (2012)	"	Cord plasma chlorpyrifos (pg/g, natural log scale) with interaction terms	"	Model 2a beta = –0.553 (− 1.943, 0.836) Model 2a interaction beta = − 1.714 (− 3.753, 0.326) for chlorpyrifos × child sex Model 2b beta = − 1.248 (− 2.270, 0.227) Model 2c beta = − 1.354 (− 2.369, − 0.339) Model 2c interaction beta = 0.024 (− 0.690, 0.738) for chlorpyrifos × parental nurturance	Family income, maternal education, child sex, and parental nurturance composite score plus: Model 2a: chlorpyrifos × child sex interaction Model 2b: parental nurturance × child sex interaction Model 2c: chlorpyrifos × parental nurturance interaction	–
Rauh et al. (2012)	Overall brain size (cm^3) at 5.9–11.2 years	Cord plasma chlorpyrifos ≥4.39 vs. <4.39 pg/g	20 ≥ 4.39 pg/g 20 < 4.39 pg/g	Mean ± SD = 1,265.1 ± 17.7 Mean ± SD = 1,242.1 ± 16.8 $P = 0.37$	Age, sex, and height	Cutoff for cord plasma chlorpyrifos was set at the highest tertile (4.39 pg/g) P-values were corrected for multiple comparisons using a false discovery rate $P < 0.05$
Rauh et al. (2012)	Morphology of cerebral surface (enlargement) at 5.9–11.2 years	"	"	Significant enlargement, especiallly of white matter, of superior temporal, posterior middle temporal, and inferior postcentral gyri bilaterally; supramarginal gyrus and inferior pariental lobule of right hemisphere; supramarginal gyrus and inferior parietal lobule of right hemisphere; and superior frontal gyrus, gyrus rectus, cuneus, and precuneus along mesial wall of right hemisphere in those with ≥ 4.39 vs. < 4.39 pg/g Significant positive dose-response relationship between chlorpyrifos and enlargement of mesial surface of superior frontal gyrus bilaterally among those with ≥ 4.39 pg/g	Age and sex, with or without overall brain size	–

Rauh et al. (2012)	"	Cord plasma chlorpyrifos ≥4.39 vs. <4.39 pg/g with interaction terms with full-scale intelligence quotient	"	Significant interaction between chlorpyrifos and intelligence quotient on surface measures in superior temporal, inferior frontal, inferior precentral, and inferior postcentral gyri bilaterally, and precuneus of left hemisphere, with positive correlation with intelligence quotient among those with <4.39 pg/g but no correlation among those with ≥4.39 pg/g Significant interaction between chlorpyrifos and intelligence quotient on surface measures in right fusiform gyrus, with inverse correlation with intelligence quotient among those with <4.39 pg/g but positive correlation among those with ≥4.39 pg/g	Age and sex	–
Rauh et al. (2012)	"	Cord plasma chlorpyrifos ≥4.39 vs. <4.39 pg/g with interaction terms with sex	19 ≥4.39 pg/g 18 <4.39 pg/g	Significant interaction between chlorpyrifos and sex on surface measures in right inferior parietal lobule, right superior marginal gyrus, and right mesial superior frontal gyrus, "reflecting disruption of normal, female-larger-than-male sex differences in the right parietal lobe and a reversal of normal, male-larger-than-female differences in the right mesial superior frontal gyrus" Significant interaction between chlorpyrifos and sex on surface measures in right dorsal parietal lobe, with positive correlation with chlorpyrifos in girls but inverse correlation in boys	"	–
Rauh et al. (2012)	Morphology of cerebral surface (deformation) at 5.9–11.2 years	Cord plasma chlorpyrifos ≥4.39 vs. <4.39 pg/g	20 ≥4.39 pg/g 20 <4.39 pg/g	Inward deformations in dorsal and mesial surfaces of left superior frontal gyrus in group with ≥4.39 pg/g	"	–
Rauh et al. (2012)	Cortical thickness at 5.9–11.2 years	"	"	"Scattered reductions" in cortical thickness in dorsal parietal and frontal cortices in group with ≥4.39 vs. <4.39 pg/g Inverse dose-response relationship between chlorpyrifos and cortical thickness in frontal pole, dorsal parietal, and orbitofrontal cortices in those with ≥4.39 pg/g	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Engel et al. (2007)	Brazelton habituation cluster before hospital discharge	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) or MDA (detectable vs. nondetectable)	144 with DAPs 153 with DMPs 144 with DEPs 148 with MDA	Beta = 0.168 (–0.230, 0.566) Beta = –0.024 (–0.335, 0.288) Beta = 0.08 (–0.300, 0.460) Beta = 0.440 (–0.145, 1.025)	Drug use during pregnancy, examiner, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Habituation = ability to respond to and inhibit discrete stimuli while asleep No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton orientation cluster before hospital discharge	"	233 with DAPs 244 with DMPs 233 with DEPs 240 with MDA	Beta = –0.106 (–0.414, 0.201) Beta = 0.018 (–0.249, 0.285) Beta = –0.028 (–0.336, 0.279) Beta = –0.100 (–0.597, 0.405)	Pre-pregnancy body mass index, examiner, neonatal jaundice, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Orientation = attention to visual and auditory stimuli and quality of overall alertness No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton motor cluster before hospital discharge	"	249 with DAPs 260 with DMPs 249 with DEPs 257 with MDA	Beta = 0.049 (–0.077, 0.174) Beta = 0.039 (–0.068, 0.146) Beta = 0.048 (–0.078, 0.174) Beta = –0.050 (–0.233, 0.156)	Infant age at examination, caffeine consumption during pregnancy, drug use during pregnancy, examiner, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Motor = motor performance and equality of movement and tone No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton range of state cluster before hospital discharge	"	253 with DAPs 264 with DMPs 253 with DEPs 256 with MDA	Beta = 0.035 (–0.120, 0.189) Beta = 0.035 (–0.096, 0.167) Beta = 0.015 (–0.140, 0.169) Beta = –0.040 (–0.281, 0.199)	Infant age at examination, examiner, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Range of state = measure of infant arousal and state lability No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton regulation of state cluster before hospital discharge	"	253 with DAPs 264 with DMPs 253 with DEPs 256 with MDA	Beta = –0.047 (–0.300, 0.207) Beta = –0.072 (–0.283, 0.138) Beta = –0.026 (–0.279, 0.227) Beta = –0.090 (–0.480, 0.303)	Maternal education, examiner, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Regulation of state = ability to regulate state in the face of increasing levels of stimulation No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton autonomic stability cluster before hospital discharge	"	253 with DAPs 264 with DMPs 253 with DEPs 256 with MDA	Beta = –0.154 (–0.382, 0.075) Beta = 0.000 (–0.192, 0.193) Beta = –0.106 (–0.334, 0.122) Beta = 0.090 (–0.274, 0.463)	Infant age at examination, examiner, smoking during pregnancy, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	Autonomic stability = signs of stress related to homeostatic adjustments of the central nervous system No significant associations with DAPs, DMPs, or DEPs categorized by quartile
Engel et al. (2007)	Brazelton number of abnormal reflexes before hospital discharge	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale or by quartile) or MDA (detectable vs. nondetectable)	239 with DAPs 250 with DMPs 239 with DEPs 242 with MDA DAPs quartile 2 DAPs quartile 3 DAPs quartile 4 DMPs quartile 2 DMPs quartile 3 DMPs quartile 4 DEPs quartile 2 DEPs quartile 3 DEPs quartile 4	Relative risk = 1.49 (1.12, 1.98) Relative risk = 1.13 (0.90, 1.41) Relative risk = 1.32 (0.99, 1.77) Relative risk = 2.24 (1.55, 3.24) Relative risk = 1.91 (1.12, 3.28) Relative risk = 1.22 (0.70, 2.11) Relative risk = 1.58 (0.96, 2.58) Relative risk = 1.58 (0.94, 2.65) Relative risk = 1.46 (0.83, 2.54) Relative risk = 1.62 (0.98, 2.66) Relative risk = 1.29 (0.71, 2.33) Relative risk = 2.59 (1.54, 4.35) Relative risk = 1.53 (0.88, 2.66)	Examiner, anesthesia during delivery, PON1 enzyme activity, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	In exploratory analyses of specific abnormal reflexes, detectable MDA levels were significantly associated with abnormal "crawling" and "resist arms" reflexes, and higher DEP levels were associated with an abnormal "crawling" reflex

Engel et al. (2007)	Brazelton ≥ 2 abnormal reflexes before hospital discharge	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) or MDA (detectable vs. nondetectable)	120 with DAPs at age 1 day 118 with DAPs at age 2 + days 126 with DMPs at age 1 day 123 with DMPs at age 2 + days 129 with DEPs at age 1 day 118 with DEPs at age 2 + days 120 with MDA at age 1 day 121 with MDA at age 2 + days	Relative risk = 1.15 (0.80, 1.63) Relative risk = 1.69 (1.11, 2.59) P-interaction > 0.10 by age Relative risk = 1.00 (0.75, 1.32) Relative risk = 1.44 (1.02, 2.03) P-interaction ≤ 0.10 by age Relative risk = 1.39 (0.96, 2.01) Relative risk = 1.60 (0.98, 2.60) P-interaction > 0.10 by age Relative risk = 2.51 (1.61, 3.90) Relative risk = 1.34 (0.72, 2.49) P-interaction ≤ 0.10 by age	Examiner, anesthesia during delivery, PON1 enzyme activity, and urinary crevaatinine; all models other than for DEPs also adjusted for overdispersion	–
Engel et al. (2007)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by maternal PON1 expression level	NR	DAPs, low PON1: relative risk = 2.38 (1.37, 4.15) DAPs, medium PON1: relative risk = 1.75 (0.96, 3.17) DAPs, high PON1: relative risk = 0.76 (0.48, 1.20) P-interaction of low and medium vs. high PON1 = < 0.05 and ≥ 0.05 DMPs, low PON1: relative risk = 1.96 (1.27, 3.03) DMPs, medium PON1: relative risk = 1.66 (1.03, 2.65) DMPs, high PON1: relative risk = 0.73 (0.56, 0.96) P-interaction of low and medium vs. high PON1 = 0.002 and 0.001 DEPs, low PON1: relative risk = 1.78 (1.01, 3.14) DEPs, medium PON1: relative risk = 1.42 (0.85, 2.35) DEPs, high PON1: relative risk = 1.56 (1.01, 2.39) P-interaction ≥ 0.05	Examiner, anesthesia during delivery, and urinary creatinine; all models other than for DEPs also adjusted for overdispersion	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Engel et al. (2011)	Bayley Mental Development Index at 12 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/Hispanics 38 whites	Tertile 1 adj. mean, total = 97.0 (93.7, 100.3) Tertile 2 adj. mean, total = 95.8 (92.5, 99.1) Tertile 3 adj. mean, total = 96.1 (93.1, 99.0) Beta, total = –1.00 (–3.28, 1.28) Tertile 1 adj. mean, blacks/Hispanics = 96.2 (92.9, 99.4) Tertile 2 adj. mean, blacks/Hispanics = 94.4 (91.2, 97.5) Tertile 3 adj. mean, blacks/Hispanics = 91.5 (88.3, 94.7) Beta, blacks/Hispanics = –3.29 (–5.88, –0.70) Tertile 1 adj. mean, whites = 92.0 (85.4, 98.7) Tertile 2 adj. mean, whites = 95.9 (90.6, 101.3) Tertile 3 adj. mean, whites = 103.7 (98.5, 108.8) P-interaction by race < 0.001 Beta, whites = 4.77 (0.69, 8.86) P-interaction by race = 0.001	Maternal age at enrollment, child sex, examiner, maternal PON1 enzyme activity, season of urine collection, laboratory batch, Home Observation for Measurement of the Environment score, alcohol consumption during pregnancy, urinary creatinine, and race/ethnicity (if not stratified; also adjusted for biomarker × race/ethnicity if stratified)	Mental Development Index rates cognitive ability in areas including memory, habituation, problem-solving, early number concepts, generalization, classification, vocalizations, language, and social skills; age-standardized to mean of 100 and SD of 15 Distinct patterns by race/ethnicity at 12 months were also observed by public vs. private housing (results NR) No significant interactions ($P \geq 0.20$) were detected between metabolite levels and *PON1* L55M or –108C > T polymorphisms or with PON1 enzyme activity on neurodevelopment at any age (data not shown)
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/Hispanics 38 whites	Tertile 1 adj. mean, total = 96.8 (93.5–100.0) Tertile 2 adj. mean, total = 96.1 (92.9–99.3) Tertile 3 adj. mean, total = 96.1 (93.4–99.0) Beta, total = –1.12 (–3.14–0.89) Tertile 1 adj. mean, blacks/Hispanics = 96.3 (93.0–99.5) Tertile 2 adj. mean, blacks/Hispanics = 94.2 (91.0–97.4) Tertile 3 adj. mean, blacks/Hispanics = 92.1 (89.0–95.2) Beta, blacks/Hispanics = –3.35 (–5.64 to –1.06) Tertile 1 adj. mean, whites = 92.2 (85.6–98.7) Tertile 2 adj. mean, whites = 97.2 (91.1–102.6) Tertile 3 adj. mean, whites = 103.3 (97.9–108.7) P-interaction by race < 0.01 Beta, whites = 4.45 (0.82–8.08) P-interaction by race < 0.001	"	—

Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/ Hispanics 38 whites	Tertile 1 adj. mean, total = 95.9 (92.9. 98.9) Tertile 2 adj. mean, total = 95.4 (92.3, 98.6) Tertile 3 adj. mean, total = 97.5 (94.3, 100.6) Beta, total = 0.03 (–2.23, 2.29) Tertile 1 adj. mean, blacks/ Hispanics = 94.3 (90.9, 97.6) Tertile 2 adj. mean, blacks/ Hispanics = 93.8 (90.4, 97.1) Tertile 3 adj. mean, blacks/ Hispanics = 95.2 (91.9, 98.6) Beta, blacks/Hispanics = –0.33 (–3.00, 2.35) Tertile 1 adj. mean, whites = 97.3 (91.8, 102.7) Tertile 2 adj. mean, whites = 96.8 (90.8, 102.9) Tertile 3 adj. mean, whites = 100.6 (94.6, 106.5) P-interaction by race = 0.82 Beta, whites = 0.86 (–3.16, 4.87) P-interaction by race = 0.62	"	–
Engel et al. (2011)	Bayley Mental Development Index at 24 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	208	Tertile 1 adj. mean = 93.6 (89.1, 98.0) Tertile 2 adj. mean = 90.8 (86.3, 95.3) Tertile 3 adj. mean = 90.3 (85.9, 94.7) Beta = –2.08 (–4.60, 0.44)	Maternal age at enrollment, child sex, examiner, maternal education, maternal PON1 enzyme activity, season of urine collection, laboratory batch, Home Observation for Measurement of the Environment score, alcohol consumption during pregnancy, urinary creatinine, and race/ethnicity	Results were not heterogeneous by exact age at 24-month testing (results NR)
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	208	Tertile 1 adj. mean = 92.5 (88.0, 96.9) Tertile 2 adj. mean = 92.9 (88.6, 97.1) Tertile 3 adj. mean = 91.1 (86.9, 95.3) Beta = –0.93 (–3.11, 1.25)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	208	Tertile 1 adj. mean = 92.6 (88.2, 97.0) Tertile 2 adj. mean = 90.5 (86.1, 94.9) Tertile 3 adj. mean = 91.2 (86.6, 95.7) Beta = –1.47 (–3.99, 1.04)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Engel et al. (2011)	Bayley Mental Development Index at 12 or 24 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by maternal $PON1_{192}$ genotype	28 blacks/Hispanics with $PON1_{192}$ QQ, 12 months	Beta = 5.72 (–0.48, 11.92)	Maternal age at enrollment, child sex, examiner, Home Observation for Measurement of the Environment score, alcohol consumption during pregnancy, laboratory batch, season of urine collection, urinary creatinine, and biomarker × genotype interaction; 24-month model also adjusted for maternal race/ethnicity	Interactions between DAPs, DMPs, and DEPs and $PON1_{192}$ genotype were detected among blacks and Hispanics at 12 months, but not at 24 months (results NR)
			82 blacks/Hispanics with $PON1_{192}$ QR/RR, 12 months	Beta = –4.94 (–7.81, –2.07)		
				P-interaction by genotype < 0.01		Results were similar when stratified by child genotype (available for 57% of subjects; results NR)
			57 all races with $PON1_{192}$ QQ, 24 months	Beta = –1.04 (–6.06, 3.99)		
			140 all races with $PON1_{192}$ QR/RR, 24 months	Beta = –1.27 (–4.40, 1.84)		
				P-interaction by genotype = 0.93		
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by maternal $PON1_{192}$ genotype	28 blacks/Hispanics with $PON1_{192}$ QQ, 12 months	Beta = 3.69 (–0.97, 8.36)	"	–
			82 blacks/Hispanics with $PON1_{192}$ QR/RR, 12 months	Beta = –1.95 (–5.36, 1.47)		
				P-interaction by genotype = 0.06		
			57 all races with $PON1_{192}$ QQ, 24 months	Beta = –0.55 (–4.79, 3.70)		
			140 all races with $PON1_{192}$ QR/RR, 24 months	Beta = –0.15 (–3.51, 3.21)		
				P-interaction by genotype = 0.88		
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by maternal $PON1_{192}$ genotype	28 blacks/Hispanics with $PON1_{192}$ QQ, 12 months	Beta = 2.76 (–2.44, 7.97)	"	–
			82 blacks/Hispanics with $PON1_{192}$ QR/RR, 12 months	Beta = –4.47 (–7.05, –1.89)		
				P-interaction by genotype = 0.0.02		
			57 all races with $PON1_{192}$ QQ, 24 months	Beta = 0.12 (–4.17, 4.42)		
			140 all races with $PON1_{192}$ QR/RR, 24 months	Beta = –0.48 (–3.27, 2.30)		
				P-interaction by genotype = 0.81		

Engel et al. (2011)	Bayley Psychomotor Development Index at 12 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/ Hispanics 38 whites	Tertile 1 adj. mean, total = 95.3 (90.9, 99.8) Tertile 2 adj. mean, total = 96.6 (92.1, 101.1) Tertile 3 adj. mean, total = 92.5 (88.5, 96.6) Beta, total = –0.52 (–3.66. 2.62) Tertile 1 adj. mean, blacks/ Hispanics = 97.7 (93.1, 102.4) Tertile 2 adj. mean, blacks/ Hispanics = 97.5 (93.0, 102.1) Tertile 3 adj. mean, blacks/ Hispanics = 94.2 (89.5, 98.9) Beta, blacks/Hispanics = –1.52 (–5.21, 2.16) Tertile 1 adj. mean, whites = 90.0 (80.5, 99.6) Tertile 2 adj. mean, whites = 97.0 (89.2, 104.7) Tertile 3 adj. mean, whites = 90.8 (83.3, 98.2) P-interaction by race = 0.65 Beta, whites = 2.07 (–3.83, 7.96) P-interaction by race = 0.31	Maternal age at enrollment, child sex, examiner, maternal PON1 enzyme activity, season of urine collection, laboratory batch, Home Observation for Measurement of the Environment score, alcohol consumption during pregnancy, urinary creatinine, and race/ethnicity (if not stratified; also adjusted for biomarker × race/ ethnicity if stratified)	Psychomotor Development Index rates fine and gross motor coordination; age-standardized to a mean of 100 and SD of 15 Metabolites were not associated with Psychomotor Development Index at 24 months (results NR)
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/Hispanics 38 whites	Tertile 1 adj. mean, total = 95.3 (91.2, 99.5) Tertile 2 adj. mean, total = 94.5 (90.1, 98.9) Tertile 3 adj. mean, total = 93.6 (89.3, 98.0) Beta, total = –0.20 (–3.28, 2.87) Tertile 1 adj. mean, blacks/ Hispanics = 97.7 (93.1, 102.4) Tertile 2 adj. mean, blacks/ Hispanics = 95.9 (91.2, 100.6) Tertile 3 adj. mean, blacks/ Hispanics = 95.6 (91.0, 100.2) Beta, blacks/Hispanics = –0.48 (–4.11, 3.16) Tertile 1 adj. mean, whites = 92.1 (84.6, 99.6) Tertile 2 adj. mean, whites = 94.4 (86.0, 102.7) Tertile 3 adj. mean, whites = 91.7 (83.5, 99.9) P-interaction by race = 0.25 Beta, whites = 0.46 (–5.12, 603) P-interaction by race = 0.78	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	149 total 111 blacks/Hispanics 38 whites	Tertile 1 adj. mean, total = 95.1 (90.7, 99.5) Tertile 2 adj. mean, total = 93.7 (89.3, 98.0) Tertile 3 adj. mean, total = 94.5 (90.6, 98.5) Beta, total = –0.92 (–3.68, 1.85) Tertile 1 adj. mean, blacks/Hispanics = 97.8 (93.2, 102.4) Tertile 2 adj. mean, blacks/Hispanics = 94.5 (90.1, 99.0) Tertile 3 adj. mean, blacks/Hispanics = 96.4 (92.0, 100.8) Beta, blacks/Hispanics = –1.81 (–5.07, 1.45) Tertile 1 adj. mean, whites = 92.5 (84.9, 100.2) Tertile 2 adj. mean, whites = 94.4 (86.7, 102.1) Tertile 3 adj. mean, whites = 89.5 (80.2, 98.8) P-interaction by race = 0.83 Beta, whites = 1.36 (–3.83, 6.56) P-interaction by race = 0.31	"	–
Engel et al. (2011)	Bayley Psychomotor Development Index at 24 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	210	Tertile 1 adj. mean = 94.8 (90.5, 99.1) Tertile 2 adj. mean = 94.5 (90.2, 98.8) Tertile 3 adj. mean = 95.1 (90.9, 99.2) Beta = 0.93 (–1.41, 3.28)	Maternal age at enrollment, child sex, examiner, maternal education, maternal PON1 enzyme activity, season of urine collection, laboratory batch, Home Observation for Measurement of the Environment score, alcohol consumption during pregnancy, urinary creatinine, and race/ethnicity	Results were not heterogeneous by exact age at 24-month testing (results NR)
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	210	Tertile 1 adj. mean = 94.7 (90.5, 98.9) Tertile 2 adj. mean = 94.9 (90.6, 99.1) Tertile 3 adj. mean = 94.8 (90.5, 99.1) Beta = 0.36 (–1.70, 2.43)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	210	Tertile 1 adj. mean = 94.8 (90.5, 99.0) Tertile 2 adj. mean = 95.4 (91.4, 99.5) Tertile 3 adj. mean = 94.2 (90.2, 98.1) Beta = 0.67 (–1.72, 3.06)	"	–

Engel et al. (2011)	Wechsler full-scale intelligence quotient at 6–9 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –1.39 (–4.54, 1.77) Beta = –1.10 (–5.01, 2.81) Beta = –1.14 (–4.55, 2.28)	Sex, race/ethnicity, maternal education, language in the home, alcohol consumption during pregnancy, laboratory batch, season of urine collection, urinary creatinine, Wechsler version (if combined), and maternal PON1 enzyme activity (unless stratified by genotype)	Subtests of Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition: Block Design, Information, Matrix Reasoning, Vocabulary, Picture Concepts, Symbol Search, Word Reasoning, and Coding Subtests of Wechsler Intelligence Scale for Children, 4th Edition: Block Design, Similarities, Digit Span, Picture Concepts, Coding, Vocabulary, Letter-Number Sequence, Matrix Reasoning, Comprehension, and Symbol Search Associations with Wechsler outcomes were not heterogeneous by race/ethnicity (results NR)
Engel et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –0.66 (–4.33, 3.00) Beta = –2.33 (–8.40, 3.74) P-interaction = 0.64	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	142 ages 6–9 years 115 ages 7–9 years 98 age 6 years	Beta = –0.46 (–3.17, 2.26) Beta = –0.39 (–3.64, 2.86) Beta = –0.56 (–3.68, 2.56)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = 0.28 (–2.89, 3.44) Beta = –1.79 (–6.83, 3.25) P-interaction = 0.49	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –2.89 (–6.15, 0.36) Beta = –3.15 (–7.19, 0.89) Beta = –1.40 (–5.27, 2.47)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –2.32 (–6.49, 1.86) Beta = –3.13 (–8.21, 1.96) P-interaction = 0.80	"	–
Engel et al. (2011)	Wechsler perceptual reasoning at 6–9 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –2.36 (–6.04, 1.31) Beta = –2.39 (–6.97, 2.19) Beta = –2.07 (–5.66, 1.52)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –0.56 (–4.80, 3.68) Beta = –7.52 (–14.53, –0.51) P-interaction = 0.09	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	142 ages 6–9 years 115 ages 7–9 years 98 age 6 years	Beta = –1.15 (–4.31, 2.02) Beta = –1.24 (–5.05, 2.57) Beta = –1.46 (–4.74, 1.83)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = 0.71 (–2.96, 4.38) Beta = –6.15 (–11.99, –0.31) P-interaction = 0.05	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –3.51 (–7.31, 0.30) Beta = –4.37 (–9.10, 0.36) Beta = –1.59 (–5.68, 2.50)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –3.24 (–8.11, 1.62) Beta = –4.80 (–10.73, 1.13) P-interaction = 0.68	"	–
Engel et al. (2011)	Wechsler verbal comprehension at 6–9 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –0.42 (–3.45, 2.62) Beta = 0.56 (–3.11, 4.23) Beta = –1.16 (–4.59, 2.27)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –0.33 (–3.87, 3.20) Beta = 0.73 (–5.12, 6.59) P-interaction = 0.76	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	142 ages 6–9 years 115 ages 7–9 years 98 age 6 years	Beta = –0.05 (–2.64, 2.54) Beta = 0.39 (–2.65, 3.42) Beta = –0.52 (–3.67, 2.62)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = 0.12 (–2.93, 3.16) Beta = 0.24 (–4.60, 5.09) P-interaction = 0.97	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	140 ages 6–9 years 114 ages 7–9 years 96 age 6 years	Beta = –1.20 (–4.35, 1.96) Beta = –0.08 (–3.91, 3.76) Beta = –2.27 (–6.14, 1.60)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by $PON1_{192}$ genotype	101 PON_{192} QR/RR 39 PON_{192} QQ	Beta = –0.45 (–4.51, 3.60) Beta = –1.20 (–6.13, 3.74) P-interaction = 0.81	"	–
Engel et al. (2011)	Wechsler processing speed at 6–9 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	114 ages 7–9 years 96 age 6 years	Beta = –1.05 (–5.57, 3.46) Beta = –1.22 (–5.12, 2.67)	"	–

Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	115 ages 7–9 years 98 age 6 years	Beta = –0.79 (–4.52, 2.94) Beta = –0.84 (–4.35, 2.67)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	114 ages 7–9 years 96 age 6 years	Beta = –2.11 (–6.81, 2.59) Beta = –1.85 (–6.25, 2.56)	"	–
Engel et al. (2011)	Wechsler working memory at 7–9 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	114	Beta = –0.53 (–4.24, 3.18)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	115	Beta = 0.29 (–2.81, 3.38)	"	–
Engel et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	114	Beta = –3.48 (–7.29, 0.34)	"	–
Young et al. (2005)	Brazelton habituation cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	175 total 107 age ≤ 3 days 66 age > 3 days	Beta = 0.03 (–0.34, 0.40) Beta, age ≤ 3 days = 0.10 (–0.40, 0.60) Beta, age > 3 days = 0.06 (–0.54, 0.66) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, smoking, alcohol, method of delivery, minutes since fed at assessment, and interviewer No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Habituation cluster includes light, rattle, bell, and pin-prick Median age at assessment: 3 days (IQR: 1–26) "Urinary metabolite levels measured at the two points during pregnancy were not significantly correlated with each other or with the post-delivery measurement, with all estimated correlations below 0.1 for total DAP, dimethyl, and diethylphosphate metabolite levels."
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	175 total 107 age ≤ 3 days 66 age > 3 days	Beta = –0.06 (–0.39, 0.27) Beta, age ≤ 3 days = –0.04 (–0.49, 0.40) Beta, age > 3 days = 0.04 (–0.50, 0.58) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	175 total 107 age ≤ 3 days 66 age > 3 days	Beta = 0.33 (–0.06, 0.72) Beta, age ≤ 3 days = 0.47 (–0.05, 0.99) Beta, age > 3 days = 0.20 (–0.43, 0.83) No association with maternal post-delivery urinary metabolites (results NR)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Young et al. (2005)	Brazelton orientation cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	379 total 197 age ≤ 3 days 182 age > 3 days	Beta = –0.17 (–0.50, 0.17) Beta, age ≤ 3 days = –0.02 (–0.53, 0.49) Beta, age > 3 days = –0.13 (–0.54, 0.27) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, interviewer, and number of prenatal care visits No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Orientation cluster includes inanimate visual, inanimate auditory, inanimate visual-auditory, animate visual, animate auditory, animate visual-auditory, and alertness
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	379 total 197 age ≤ 3 days 182 age > 3 days	Beta = –0.12 (–0.43, 0.19) Beta, age ≤ 3 days = –0.08 (–0.54, 0.39) Beta, age > 3 days = 0.01 (–0.37, 0.38) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	379 total 197 age ≤ 3 days 182 age > 3 days	Beta = –0.32 (–0.66, 0.03) Beta, age ≤ 3 days = –0.11 (–0.65, 0.43) Beta, age > 3 days = –0.33 (–0.73, 0.08) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	Brazelton motor performance cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.03 (–0.19, 0.14) Beta, age ≤ 3 days = 0.04 (–0.20, 0.28) Beta, age > 3 days = –0.07 (–0.28, 0.15) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, poverty level, gestational age at initiation of prenatal care, and interviewer No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Motor performance cluster includes tonus, maturity, pull-to-sit, defense, and activity
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.05 (–0.20, 0.10) Beta, age ≤ 3 days = 0.03 (–0.19, 0.24) Beta, age > 3 days = –0.11 (–0.31, 0.09) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.10 (–0.06, 0.27) Beta, age ≤ 3 days = 0.08 (–0.17, 0.33) Beta, age > 3 days = 0.17 (–0.05, 0.38) No association with maternal post-delivery urinary metabolites (results NR)	"	–

Young et al. (2005)	Brazelton range of state cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.09 (–0.16, 0.34) Beta, age ≤ 3 days = 0.11 (–0.21, 0.43) Beta, age > 3 days = –0.02 (–0.44, 0.40) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, number of prenatal care visits, gestational age at initiation of prenatal care, alcohol, and interviewer No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Range of state cluster includes peak of excitement, rapidity of build-up, irritability, and lability of state
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.08 (–0.15, 0.32) Beta, age ≤ 3 days = 0.17 (–0.12, 0.46) Beta, age > 3 days = –0.12 (–0.51, 0.27) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.02 (–0.27, 0.24) Beta, age ≤ 3 days = –0.21 (–0.54, 0.12) Beta, age > 3 days = 0.20 (–0.21, 0.62) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	Brazelton regulation of state cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.07 (–0.39, 0.24) Beta, age ≤ 3 days = –0.07 (–0.50, 0.36) Beta, age > 3 days = –0.10 (–0.58, 0.37) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, pre-pregnancy body mass index, infant sex, parity, caffeine use, and interviewer No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Regulation of state cluster includes cuddliness, consolability, self-quieting, and hand-to-mouth
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.05 (–0.34, 0.24) Beta, age ≤ 3 days = –0.06 (–0.45, 0.33) Beta, age > 3 days = –0.06 (–0.50, 0.39) No association with maternal post-delivery urinary metabolites (results NR)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.15 (–0.47, 0.17) Beta, age ≤ 3 days = –0.08 (–0.52, 0.37) Beta, age > 3 days = –0.24 (–0.72, 0.24) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	Brazelton autonomic stability cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.16 (–0.36, 0.05) Beta, age ≤ 3 days = –0.09 (–0.38, 0.20) Beta, age > 3 days = –0.19 (–0.49, 0.12) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, infant sex, parity, vitamin use, minutes since fed at assessment, interviewer, and illicit drug use during pregnancy No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Autonomic stability cluster includes tremors, startles, and skin color
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = –0.17 (–0.35, 0.02) Beta, age ≤ 3 days = –0.15 (–0.42, 0.11) Beta, age > 3 days = –0.14 (–0.43, 0.14) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.06 (–0.15, 0.27) Beta, age ≤ 3 days = 0.31 (0.01, 0.61) Beta, age > 3 days = –0.16 (–0.47, 0.14) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	Brazelton reflexes cluster at < 2 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.23 (0.05, 0.41) Beta, age ≤ 3 days = –0.01 (–0.24, 0.22) Beta, age > 3 days = 0.53 (0.23, 0.82) No association with maternal post-delivery urinary metabolites (results NR)	Age at assessment, maternal age at delivery, smoking, vitamin use, interviewer, and mean diastolic and systolic blood pressure No change after additional adjustment for maternal post-delivery urinary metabolite levels, birth weight, or gestational age; some attenuation after adjustment for creatinine (results NR)	Reflex cluster includes plantar, Babinski, ankle clonus, rooting, sucking, glabella, passive resistance of legs, passive resistance of arms, palmar, placing, standing, walking, crawling, incurvation, tonic deviation of head and eyes, nystagmus, tonic neck reflex, and moro reflex
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.18 (0.02, 0.34) Beta, age ≤ 3 days = –0.00 (–0.21, 0.20) Beta, age > 3 days = 0.41 (0.12, 0.69) No association with maternal post-delivery urinary metabolites (results NR)	"	–

Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, log_{10} scale)	381 total 197 age ≥ 3 days 184 age > 3 days	Beta = 0.22 (0.04, 0.40) Beta, age ≤ 3 days = 0.08 (–0.16, 0.32) Beta, age > 3 days = 0.37 (0.09, 0.64) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	Brazelton > 3 abnormal reflexes at > 3 days to < 2 months	Maternal prenatal urinary DAPs (nmol/L, log_{10} scale): 1.07–1.65 1.65–1.83 1.83–2.07 2.08–2.30 2.31–3.17	3 of 37 5 of 37 6 of 37 5 of 37 12 of 36	Proportion = 8% Proportion = 14% Proportion = 16% Proportion = 14% Proportion = 33% P-trend = 0.01 Odds ratio per unit increase = 4.9 (1.5, 16.1) No association with maternal post-delivery urinary metabolites (results NR)	NR	–
Young et al. (2005)	"	Maternal prenatal urinary DMPs (nmol/L, log_{10} scale): 0.68–1.48 1.48–1.69 1.70–1.95 1.95–2.19 2.19–3.15	3 of 37 5 of 37 6 of 37 9 of 37 8 of 36	Proportion = 8% Proportion = 14% Proportion = 16% Proportion = 24% Proportion = 22% P-trend = 0.03 Odds ratio per unit increase = 3.2 (1.1, 9.8) No association with maternal post-delivery urinary metabolites (results NR)	"	–
Young et al. (2005)	"	Maternal prenatal urinary DEPs (nmol/L, log_{10} scale): 0.51–0.90 0.90–1.10 1.11–1.27 1.28–1.58 1.58–2.35	6 of 37 2 of 37 6 of 37 5 of 37 12 of 36	Proportion = 16% Proportion = 5% Proportion = 16% Proportion = 14% Proportion = 33% P-trend = 0.05 Odds ratio per unit increase = 3.4 (1.2, 9.9) No association with maternal post-delivery urinary metabolites (results NR)	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2007)	Bayley Mental Development Index at 6, 12, or 24 months	Maternal or child urinary DAPs (nmol/L, $\log_{10}$ scale)	395 at 6 months 393 at 12 months 369 at 24 months	Prenatal beta = –1.15 (–2.89, 0.59) Child beta = –0.17 (–1.23, 0.90) Prenatal beta = –1.34 (–3.59, 0.92) Child beta = 1.36 (–0.05, 2.78) Prenatal beta = –3.54 (–6.59, –0.49) Child beta = 2.37 (0.50, 4.24)	Psychometrician, location, exact age at assessment, sex, breast-feeding duration, score on Infant-Toddler Home Observation for Measurement of the Environment instrument, household income above poverty threshold, parity, and maternal Peabody Picture Vocabulary Test score	Mean ± SD age (months) at child assessments: 6.6 ± 1.1, 12.8 ± 1.6, and 24.6 ± 1.1 Bayley Scales of Infant Development are standardized by age to a mean ± SD of 100 ± 15; scores < 85 indicate possible developmental delay Longitudinal analyses of DAPs and Bayley scores produced similar findings (not reported)
Eskenazi et al. (2007)	"	Maternal or child urinary DMPs (nmol/L, $\log_{10}$ scale)	395 at 6 months 393 at 12 months 369 at 24 months	Prenatal beta = –0.95 (–2.52, 0.62) Child beta = –0.31 (–1.28, 0.67) Prenatal beta = –1.06 (–3.12, 0.99) Child beta = 0.75 (–0.44, 1.93) Prenatal beta = –3.64 (–6.36, –0.91) Child beta = 2.01 (0.24, 3.78)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DEPs (nmol/L, $\log_{10}$ scale)	395 at 6 months 393 at 12 months 369 at 24 months	Prenatal beta = –0.16 (–1.96, 1.65) Child beta = 0.24 (–0.78, 1.25) Prenatal beta = –1.14 (–3.51, 1.22) Child beta = 1.89 (0.21, 3.58) Prenatal beta = –0.85 (–3.98, 2.27) Child beta = 1.02 (–0.52, 2.57)	"	–
Eskenazi et al. (2007)	"	Maternal urinary MDA (µg/L)	39% detectable	At 6 months: Undetectable: beta = referent Detectable < median: beta = 0.98 (–0.85, 2.81) Detectable ≥ median: beta = –0.25 (–2.10, 1.60) At 12 months: Undetectable: beta = referent Detectable < median: beta = 0.95 (–1.55, 3.46) Detectable ≥ median: beta = 2.40 (–0.13, 4.94) At 24 months: Undetectable: beta = referent Detectable < median: beta = –1.09 (–4.51, 2.32) Detectable ≥ median: beta = 0.24 (–3.03, 3.52)	"	–

Eskenazi et al. (2007)	"	Maternal urinary TCPy (μg/L)	91% detectable	At 6 months: Undetectable: beta = referent Detectable < median: beta = 0.24 (–2.12, 2.61) Detectable ≥ median: beta = 0.08 (–2.29, 2.44) At 12 months: Undetectable: beta = referent Detectable < median: beta = –0.45 (–3.67, 2.76) Detectable ≥ median: beta = –0.65 (–3.88, 2.58) At 24 months: Undetectable: beta = referent Detectable < median: beta = –1.02 (–5.34, 3.31) Detectable ≥ median: beta = –1.94 (–6.26, 2.37)	"	–
Eskenazi et al. (2007)	Bayley Psychomotor Development Index at 6, 12, or 24 months	Maternal or child urinary DAPs (nmol/L, $\log_{10}$ scale)	396 at 6 months 392 at 12 months 371 at 24 months	Prenatal beta = –0.71 (–3.28, 1.86) Child beta = 0.39 (–1.18, 1.97) Prenatal beta = –0.60 (–3.77, 2.57) Child beta = 1.22 (–0.78, 3.21) Prenatal beta = –1.28 (–4.01, 1.46) Child beta = 1.06 (–0.62, 2.74)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DMPs (nmol/L, $\log_{10}$ scale)	396 at 6 months 392 at 12 months 371 at 24 months	Prenatal beta = –0.55 (–2.88, 1.77) Child beta = 0.28 (–1.17, 1.72) Prenatal beta = –1.15 (–4.03, 1.74) Child beta = 0.46 (–1.22, 2.13) Prenatal beta = –1.24 (–3.70, 1.21) Child beta = 1.01 (–0.58, 2.60)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DEPs (nmol/L, $\log_{10}$ scale)	396 at 6 months 392 at 12 months 371 at 24 months	Prenatal beta = 0.02 (–2.63, 2.67) Child beta = 0.60 (–0.89, 2.09) Prenatal beta = 0.30 (–3.03, 3.63) Child beta = 1.91 (–0.46, 4.27) Prenatal beta = –0.86 (–3.64, 1.92) Child beta = 0.30 (–1.07, 1.67)	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2007)	"	Maternal urinary MDA (μg/L)	39% detectable	At 6 months: Undetectable: beta = referent Detectable < median: beta = 0.42 (–2.34, 3.18) Detectable ≥ median: beta = –1.45 (–4.21, 1.32) At 12 months: Undetectable: beta = referent Detectable < median: beta = –0.53 (–4.05, 3.00) Detectable ≥ median: beta = 0.75 (–2.81, 4.31) At 24 months: Undetectable: beta = referent Detectable < median: beta = –0.73 (–3.87, 2.41) Detectable ≥ median: beta = 0.33 (–2.68, 3.35)	"	–
Eskenazi et al. (2007)	"	Maternal urinary TCPy (μg/L)	91% detectable	At 6 months: Undetectable: beta = referent Detectable < median: beta = –0.56 (–4.03, 2.91) Detectable ≥ median: beta = –0.21 (–3.69, 3.27) At 12 months: Undetectable: beta = referent Detectable < median: beta = –0.70 (–5.26, 3.86) Detectable ≥ median: beta = –1.62 (–6.20, 2.96) At 24 months: Undetectable: beta = referent Detectable < median: beta = –2.65 (–6.50, 1.21) Detectable ≥ median: beta = –2.72 (–6.57, 1.12)	"	–
Eskenazi et al. (2007)	Child Behavior Checklist attention problems syndrome score at 24 months	Maternal or child urinary DAPs (nmol/L, $\log_{10}$ scale)	30 (8.4%) borderline	Prenatal odds ratio = 0.77 (0.27, 2.24) Child odds ratio = 1.41 (0.75, 2.64)	Sex, exact age at assessment, breast-feeding duration, score on Infant-Toddler Home Observation for Measurement of the Environment instrument, household income above poverty threshold, parity, maternal Peabody Picture Vocabulary Test score, and maternal depression	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile (N = 7, 2.0%)

Eskenazi et al. (2007)	"	Maternal or child urinary DMPs (nmol/L, $\log_{10}$ scale)	30 (8.4%) borderline	Prenatal odds ratio = 0.78 (0.31, 1.96) Child odds ratio = 1.54 (0.85, 2.76)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DEPs (nmol/L, $\log_{10}$ scale)	30 (8.4%) borderline	Prenatal odds ratio = 0.78 (0.26, 2.31) Child odds ratio = 1.02 (0.61, 1.71)	"	–
Eskenazi et al. (2007)	"	Maternal urinary MDA or TCPy (µg/L)	30 (8.4%) borderline	"no significant associations" (results NR)	"	–
Eskenazi et al. (2007)	Child Behavior Checklist ADHD score at 24 months	Maternal or child urinary DAPs (nmol/L, $\log_{10}$ scale)	34 (9.6%) borderline	Prenatal odds ratio = 1.34 (0.50, 3.59) Child odds ratio = 1.11 (0.61, 2.03)	"	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile (N = 10, 2.8%)
Eskenazi et al. (2007)	"	Maternal or child urinary DMPs (nmol/L, $\log_{10}$ scale)	34 (9.6%) borderline	Prenatal odds ratio = 1.27 (0.53, 3.04) Child odds ratio = 1.10 (0.63, 1.94)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DEPs (nmol/L, $\log_{10}$ scale)	34 (9.6%) borderline	Prenatal odds ratio = 0.59 (0.21, 1.68) Child odds ratio = 1.18 (0.72, 1.94)	"	–
Eskenazi et al. (2007)	"	Maternal urinary MDA or TCPy (µg/L)	34 (9.6%) borderline	"no significant associations" (results NR)	"	–
Eskenazi et al. (2007)	Child Behavior Checklist pervasive developmental disorder score at 24 months	Maternal or child urinary DAPs (nmol/L, $\log_{10}$ scale)	51 (14.4%) clinical	Prenatal odds ratio = 2.25 (0.99, 5.16) Child odds ratio = 1.71 (1.02, 2.87)	"	"Borderline" score > 93rd percentile (N = 105, 29.6%) "Clinical" score > 97th percentile
Eskenazi et al. (2007)	"	Maternal or child urinary DMPs (nmol/L, $\log_{10}$ scale)	51 (14.4%) clinical	Prenatal odds ratio = 2.19 (1.05, 4.58) Child odds ratio = 1.52 (0.94, 2.45)	"	–
Eskenazi et al. (2007)	"	Maternal or child urinary DEPs (nmol/L, $\log_{10}$ scale)	51 (14.4%) clinical	Prenatal odds ratio = 0.88 (0.37, 2.07) Child odds ratio = 1.72 (1.12, 2.64)	"	–
Eskenazi et al. (2007)	"	Maternal urinary MDA or TCPy (µg/L)	51 (14.4%) clinical	"no significant associations" (results NR)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2010)	Bayley Mental Development Index at 24 months	Maternal prenatal urinary DAPs (nmol/L, log_{10} scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = –3.2 (–9.8, 3.5) Beta = –3.7 (–8.0, 0.6) Beta = –5.5 (–11.1, 0.1) P-interaction = 0.98 Beta = –6.5 (–15.6, 2.6) Beta = –1.2 (–5.2, 2.9) Beta = –6.9 (–12.8, –0.9) P-interaction = 0.33	Age at assessment, sex, parity, breast-feeding duration, Infant-Toddler Home Observation for Measurement of the Environment score, maternal Peabody Picture Vocabulary Test score, household poverty status, psychometrician, and testing location	Interactions were statistically non-significant whether genotype was coded as a categorical or ordinal variable Associations were "similar, albeit weaker" when stratified by maternal genotype (not shown here)
Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, log_{10} scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = –2.2 (–8.0, 3.6) Beta = –3.4 (–7.4, 0.6) Beta = –5.9 (–11.1, –0.6) P-interaction = 0.91 Beta = –4.4 (–12.4, 3.6) Beta = –1.3 (–4.9, 2.4) Beta = –7.4 (–13.0, –1.9) P-interaction = 0.38	"	–
Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, log_{10} scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = –0.3 (–7.2, 6.7) Beta = –1.7 (–6.3, 3.0) Beta = –3.4 (–8.8, 2.1) P-interaction = 0.84 Beta = 1.4 (–8.4, 11.1) Beta = –1.1 (–5.2, 3.0) Beta = –2.5 (–8.7, 3.6) P-interaction = 0.47	"	–
Eskenazi et al. (2010)	"	Maternal prenatal urinary DAPs (nmol/L, log_{10} scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –5.4 (–11.9, 1.1) Beta = –4.3 (–11.6, 3.0) Beta = –1.2 (–8.7, 6.4) P-interaction = 0.89 Beta = –6.6 (–12.9, –0.2) Beta = –1.0 (–7.9, 5.9) Beta = –5.8 (–13.9, 2.2) P-interaction = 0.72	"	Associations across maternal PON1 enzyme levels and activities at delivery were "similar to those for cord blood enzyme levels" (not shown here)
Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, log_{10} scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –5.4 (–11.4, 0.5) Beta = –4.5 (–11.2, 2.3) Beta = –0.5 (–7.0, 6.1) P-interaction = 0.72 Beta = –6.7 (–12.6, –0.8) Beta = –0.9 (–7.2, 5.4) Beta = –4.2 (–11.2, 2.9) P-interaction = 0.97	"	–

Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –3.2 (–9.7, 3.3) Beta = 0.2 (–6.8, 7.2) Beta = 2.7 (–5.2, 10.5) P-interaction = 0.23 Beta = –3.0 (–10.0, 3.9) Beta = –1.7 (–7.8, 4.5) Beta = –0.5 (–8.5, 7.6) P-interaction = 0.68	"	–
Eskenazi et al. (2010)	Bayley Psychomotor Development Index at 24 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = –2.3 (–7.8, 3.3) Beta = –0.8 (–4.8, 3.3) Beta = –1.0 (–7.1, 5.1) P-interaction = 0.89 Beta = –1.7 (–8.7, 5.4) Beta = 0.1 (–3.5, 3.8) Beta = –5.1 (–11.1, 1.0) P-interaction = 0.53	"	Interactions were statistically non-significant whether genotype was coded as a categorical or ordinal variable Associations were "similar, albeit weaker" when stratified by maternal genotype (not shown here)
Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = –1.6 (–6.4, 3.3) Beta = –0.3 (–4.0, 3.4) Beta = –1.2 (–6.9, 4.4) P-interaction = 0.87 Beta = –2.1 (–8.3, 4.0) Beta = 0.7 (–2.6, 4.0) Beta = –4.7 (–10.4, 1.0) P-interaction = 0.36	"	–
Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Beta = 0.9 (–4.9, 6.8) Beta = –2.2 (–6.5, 2.1) Beta = –1.5 (–7.3, 4.2) P-interaction = 0.66 Beta = 4.5 (–2.9, 11.9) Beta = –1.9 (–5.6, 1.8) Beta = –3.8 (–9.9, 2.3) P-interaction = 0.14	"	–
Eskenazi et al. (2010)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –2.4 (–8.3, 3.4) Beta = –1.8 (–8.3, 4.6) Beta = 1.2 (–5.7, 8.1) P-interaction = 0.46 Beta = –4.7 (–10.6, 1.3) Beta = 0.0 (–6.6, 6.7) Beta = 1.5 (–5.4, 8.4) P-interaction = 0.69	"	Associations across maternal PON1 enzyme levels and activities at delivery were "similar to those for cord blood enzyme levels" (not shown here)
Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –1.1 (–6.5, 4.3) Beta = –3.2 (–9.2, 2.8) Beta = 1.7 (–4.3, 7.7) P-interaction = 0.41 Beta = –4.1 (–9.6, 1.4) Beta = 0.4 (–5.7, 6.5) Beta = 1.4 (–4.6, 7.4) P-interaction = 0.59	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Beta = –5.0 (–10.7, 0.7) Beta = 1.6 (–4.5, 7.8) Beta = 1.6 (–5.6, 8.8) P-interaction = 0.42 Beta = –4.7 (–11.1, 1.6) Beta = –2.5 (–8.3, 3.4) Beta = 3.7 (–3.1, 10.5) P-interaction = 0.35	"	—
Eskenazi et al. (2010)	Child Behavior Checklist pervasive developmental disorder score at 24 months	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	O = 4.2 (0.5, 36.8) Odds ratio = 2.0 (0.6, 6.0) Odds ratio = 1.9 (0.3, 10.4) P-interaction = 0.91 Odds ratio = 5.4 (0.7, 44.0) Odds ratio = 1.2 (0.4, 3.6) Odds ratio = 5.2 (0.8, 35.1) P-interaction = 0.29	Age at assessment, sex, parity, breast-feeding duration, Infant-Toddler Home Observation for Measurement of the Environment score, maternal Peabody Picture Vocabulary Test score, household poverty status, and maternal depression	Interactions were statistically non-significant whether genotype was coded as a categorical or ordinal variable Associations were "similar, albeit weaker" when stratified by maternal genotype (not shown here)
Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Odds ratio = 3.3 (0.5, 21.3) Odds ratio = 2.2 (0.8, 5.9) Odds ratio = 1.9 (0.4, 9.8) P-interaction = 0.94 Odds ratio = 4.8 (0.8, 31.1) Odds ratio = 1.2 (0.5, 3.3) Odds ratio = 6.1 (1.0, 39.3) P-interaction = 0.20	"	—
Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale) by child genotype	111 *PON1*$_{-108}$ CC 179 *PON1*$_{-108}$ CT 74 *PON1*$_{-108}$ TT 94 *PON1*$_{192}$ RR 188 *PON1*$_{192}$ QR 86 *PON1*$_{192}$ QQ	Odds ratio = 7.4 (0.6, 93.9) Odds ratio = 0.8 (0.2, 2.8) Odds ratio = 0.8 (0.1, 4.3) P-interaction = 0.44 Odds ratio = 1.0 (0.1, 8.2) Odds ratio = 0.8 (0.2, 2.6) Odds ratio = 1.2 (0.2, 7.7) P-interaction = 0.97	"	—
Eskenazi et al. (2010)	"	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	 Odds ratio = 3.7 (0.5, 30.8) Odds ratio = 2.5 (0.3, 25.0) Odds ratio = 1.8 (0.3, 11.8) P-interaction = 0.48 Odds ratio = 13.2 (1.4, 128.3) Odds ratio = 0.4 (0.0, 3.9) Odds ratio = 4.7 (0.5, 41.6) P-interaction = 0.88	"	Associations across maternal PON1 enzyme levels and activities at delivery were "similar to those for cord blood enzyme levels" (not shown here)

Eskenazi et al. (2010)	"	Maternal prenatal urinary DMPs (nmol/L, log_{10} scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	Odds ratio = 3.1 (0.5, 20.2) Odds ratio = 2.8 (0.3, 24.3) Odds ratio = 2.0 (0.4, 11.4) P-interaction = 0.43 Odds ratio = 7.3 (0.9, 56.9) Odds ratio = 0.8 (0.1, 6.2) Odds ratio = 4.5 (0.7, 30.4) P-interaction = 0.90	"	–
Eskenazi et al. (2010)	"	Maternal prenatal urinary DEPs (nmol/L, log_{10} scale) by cord blood PON1 quantity or activity	PON1 quantity: 89 tertile 1 88 tertile 2 88 tertile 3 PON1 activity: 91 tertile 1 85 tertile 2 87 tertile 3	Odds ratio = 1.7 (0.3, 11.4) Odds ratio = 1.3 (0.2, 10.6) Odds ratio = 0.9 (0.1, 8.2) P-interaction = 0.24 Odds ratio = 7.0 (0.8, 58.0) Odds ratio = 0.4 (0.1, 3.1) Odds ratio = 0.7 (0.1, 6.6) P-interaction = 0.15	"	–
Marks et al. (2010)	Child Behavior Checklist attention problems borderline at 3.5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, log_{10} scale)	17/330 (5.2%) total 12/151 (7.9%) boys 5/179 (2.8%) girls	DAPs odds ratio = 3.0 (0.7, 11.7) DAPs odds ratio, boys = 4.1 (0.8, 22.2) DAPs odds ratio, girls = 2.1 (0.2, 29.9) P-interaction by sex = 0.68 DMPs odds ratio = 3.2 (0.9, 11.3) DEPs odds ratio = 2.1 (0.6, 7.0)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, log_{10} scale)	17/289 (5.9^)	DAPs odds ratio = 1.6 (0.8, 3.5) DMPs odds ratio = 1.6 (0.8, 3.3) DEPs odds ratio = 1.9 (0.9, 3.9)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Child Behavior Checklist attention problems continuous score at 3.5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, log_{10} scale)	330 total 151 boys 179 girls	DAPs beta = 0.3 (–0.2, 0.7) DAPs beta, boys = 0.7 (0.0, 1.4) DAPs beta, girls = –0.1 (–0.7, 0.5) P-interaction by sex = 0.05 DMPs beta = 0.3 (–0.1, 0.7) DEPs beta = 0.0 (–0.5, 0.4)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	289	DAPs beta = 0.1 (–0.2, 0.4) DMPs beta = 0.1 (–0.2, 0.3) DEPs beta = 0.2 (0.0, 0.5)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Child Behavior Checklist ADHD borderline at 3.5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	18/329 (5.5%) total 12/151 (7.9%) boys 6/176 (3.4%) girls	DAPs odds ratio = 3.1 (0.8, 11.5) DAPs odds ratio, boys = 6.4 (1.1, 39.0) DAPs odds ratio, girls = 1.0 (0.1, 11.2) P-interaction by sex = 0.21 DMPs odds ratio = 1.3 (0.4, 4.4) DEPs odds ratio = 2.8 (0.9, 8.9)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	17/288 (5.9%)	DAPs odds ratio = 1.4 (0.7, 3.1) DMPs odds ratio = 1.4 (0.7, 3.0) DEPs odds ratio = 1.0 (0.5, 2.2)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Child Behavior Checklist ADHD continuous score at 3.5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	329 total 151 boys 176 girls	DAPs beta = 0.5 (–0.3, 1.3) DAPs beta, boys = 1.3 (0.1, 2.5) DAPs beta, girls = –0.2 (–1.2, 0.8) P-interaction by sex = 0.06 DMPs beta = 0.6 (–0.1, 1.3) DEPs beta = –0.2 (–0.9, 0.6)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	–
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	288	DAPs beta = 0.1 (–0.3, 0.6) DMPs beta = 0.1 (–0.3, 0.6) DEPs beta = 0.2 (–0.3, 0.7)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

Marks et al. (2010)	NEPSY-II visual attention continuous score at 3.5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	319 total 143 boys 176 girls	DAPs beta = 0.2 (–0.5, 0.8) DAPs beta, boys = 0.2 (–0.8, 1.1) DAPs beta, girls = 0.2 (–0.7, 1.2) P-interaction by sex = 0.99 DMPs beta = 0.1 (–0.5, 0.6) DEPs beta = –0.2 (–0.8, 0.5)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	NEPSY-II visual attention subtest is scaled to an age-standardized mean ± SD of 10 ± 3
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	277	DAPs beta = –0.1 (–0.5, 0.3) DMPs beta = –0.1 (–0.5, 0.3) DEPs beta = –0.1 (–0.5, 0.3)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Child Behavior Checklist attention problems borderline at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	13/322 (4.0%) total 10/154 (6.5%) boys 3/168 (1.8%) girls	DAPs odds ratio = 0.8 (0.2, 3.8) DAPs odds ratio, boys = 1.0 (0.2, 6.0) DAPs odds ratio, girls = 0.6 (0.0, 17.3) P-interaction by sex = 0.77 DMPs odds ratio = 2.0 (0.5, 8.5) DEPs odds ratio = 0.7 (0.2, 2.8)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	13/319 (4.1%)	DAPs odds ratio = 1.0 (0.4, 2.4) DMPs odds ratio = 0.9 (0.4, 2.1) DEPs odds ratio = 1.8 (0.8, 3.9)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Marks et al. (2010)	Child Behavior Checklist attention problems continuous score at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	322 total 154 boys 168 girls	DAPs beta = 0.7 (0.2, 1.2) DAPs beta, boys = 0.9 (0.2, 1.7) DAPs beta, girls = 0.4 (–0.2, 1.0) P-interaction by sex = 0.28 DMPs beta = 0.6 (0.2, 1.0) DEPs beta = 0.4 (–0.1, 0.9)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	–
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	319	DAPs beta = 0.0 (–0.3, 0.2) DMPs beta = –0.1 (–0.3, 0.2) DEPs beta = 0.0 (–0.2, 0.3)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Child Behavior Checklist ADHD borderline at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	23/322 (7.1%) total 14/154 (9.1%) boys 9/168 (5.4%) girls	DAPs odds ratio = 1.1 (0.3, 3.5) DAPs odds ratio, boys = 4.9 (0.7, 33.0) DAPs odds ratio, girls = 0.3 (0.0, 2.2) P-interaction by sex = 0.18 DMPs odds ratio = 1.3 (0.4, 4.0) DEPs odds ratio = 1.1 (0.4, 3.2)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	"Borderline" score > 93rd percentile "Clinical" score > 97th percentile
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	22/319 (6.9%)	DAPs odds ratio = 0.6 (0.3, 1.2) DMPs odds ratio = 0.5 (0.3, 1.1) DEPs odds ratio = 0.9 (0.5, 1.7)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

Marks et al. (2010)	Child Behavior Checklist ADHD continuous score at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	322 total 154 boys 168 girls	DAPs beta = 1.3 (0.4, 2.1) DAPs beta, boys = 1.9 (0.6, 3.2) DAPs beta, girls = 0.6 (–0.5, 1.6) P-interaction by sex = 0.13 DMPs beta = 1.1 (0.3, 1.9) DEPs beta = 0.7 (–0.2, 1.5)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	–
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	319	DAPs beta = 0.0 (–0.5, 0.5) DMPs beta = 0.0 (–0.5, 0.4) DEPs beta = 0.1 (–0.3, 0.6)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Conners markedly atypical % omissions at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	59/312 (18.9%) total 21/148 (14.2%) boys 38/164 (23.2%) girls	DAPs odds ratio = 1.5 (0.7, 3.3) DAPs odds ratio, boys = 1.7 (0.4, 6.4) DAPs odds ratio, girls = 1.4 (0.5, 4.0) P-interaction by sex = 0.90 DMPs odds ratio = 1.9 (0.9, 4.1) DEPs odds ratio = 1.3 (0.6, 2.8)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	Conners' Kiddie Continuous Performance Test is scaled to an age-standardized mean ± SD of 50 ± 10, with score > 65 considered "markedly atypical"
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	58/309 (18.8%)	DAPs odds ratio = 1.0 (0.6, 1.6) DMPs odds ratio = 0.9 (0.6, 1.5) DEPs odds ratio = 1.5 (1.0, 2.2)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Marks et al. (2010)	Conners markedly atypical % commissions at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	54/312 (17.3%) 24/148 (14.2%) boys 30/164 (18.3%) girls	DAPs odds ratio = 1.0 (0.5, 2.2) DAPs odds ratio, boys = 0.9 (0.2, 3.2) DAPs odds ratio, girls = 1.2 (0.4, 3.3) P-interaction by sex = 0.89 DMPs odds ratio = 1.2 (0.6, 2.7) DEPs odds ratio = 0.8 (0.4, 1.6)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score	–
					Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	53/309 (17.2%)	DAPs odds ratio = 1.1 (0.7, 1.7) DMPs odds ratio = 1.1 (0.7, 1.8) DEPs odds ratio = 0.9 (0.6, 1.4)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score	–
					Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	
Marks et al. (2010)	Conners markedly atypical hit reaction time at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	20/311 (6.4%) total 7/147 (4.8%) boys 13/164 (7.9%) girls	DAPs odds ratio = 1.6 (0.5, 5.2) DAPs odds ratio, boys = 1.2 (0.1, 11.5) DAPs odds ratio, girls = 1.7 (0.4, 7.4) P-interaction by sex = 0.72 DMPs odds ratio = 1.1 (0.3, 3.6) DEPs odds ratio = 1.5 (0.5, 4.6)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score	–
					Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	19/308 (6.2%)	DAPs odds ratio = 1.1 (0.5, 2.3) DMPs odds ratio = 1.0 (0.5, 2.0) DEPs odds ratio = 1.3 (0.7, 2.4)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score	–
					Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	

Marks et al. (2010)	ADHD Confidence Index > 70th percentile at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	25/297 (8.4%) total 14/140 (10.0%) boys 11/157 (7.0%) girls	DAPs odds ratio = 5.1 (1.7, 15.7) DAPs odds ratio, boys = 10.1 (1.6, 65.3) DAPs odds ratio, girls = 3.3 (0.6, 17.0) P-interaction by sex = 0.41 DMPs odds ratio = 6.6 (2.2, 19.3) DEPs odds ratio = 3.2 (1.2, 8.9)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	ADHD Confidence Index score on Conners' Kiddie Continuous Performance Test is scaled to a range of 0–100, with > 70th percentile considered as clinical ADHD
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	24/294 (8.2%)	DAPs odds ratio = 1.3 (0.7, 2.5) DMPs odds ratio = 1.2 (0.7, 2.3) DEPs odds ratio = 1.5 (0.8, 2.8)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	ADHD Confidence Index continuous score at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	297 total 140 boys 157 girls	DAPs beta = 3.4 (–1.8, 8.7) DAPs beta, boys = 6.3 (–0.5, 13.3) DAPs beta, girls = 0.5 (–7.2, 8.3) P-interaction by sex = 0.39 DMPs beta = 2.0 (–2.8, 6.9) DEPs beta = 3.4 (–1.7, 8.6)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	–
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	294	DAPs beta = –0.7 (–3.8, 2.3) DMPs beta = –1.0 (–3.9, 1.9) DEPs beta = 2.2 (–0.5, 5.0)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Marks et al. (2010)	Hillside Behavior Rating Scale Attention ≥ 7 of 12 at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	23/322 (7.1%) total 14/153 (9.2%) boys 9/169 (5.3%) girls	DAPs odds ratio = 3.0 (0.9, 9.8) DAPs odds ratio, boys = 7.9 (1.4, 46.0) DAPs odds ratio, girls = 1.0 (0.2, 5.9) P-interaction by sex = 0.14 DMPs odds ratio = 2.3 (0.7, 7.4) DEPs odds ratio = 2.9 (1.0, 8.5)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	Hillside Behavior Rating Scale score is scaled to a range of 0–12, with score ≥ 7 (< 10% of children) considered as displaying "a higher degree of attention problems"
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	23/319 (7.2%)	DAPs odds ratio = 1.4 (0.7, 2.8) DMPs odds ratio = 1.1 (0.6, 2.1) DEPs odds ratio = 1.4 (0.8, 2.6)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–
Marks et al. (2010)	Composite ADHD indicator at 5 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	27/319 (8.5%) total 19/150 (12.7%) boys 8/169 (4.7%) girls	DAPs odds ratio = 3.5 (1.1, 10.7) DAPs odds ratio, boys = 11.1 (1.8, 66.5) DAPs odds ratio, girls = 1.1 (0.2, 7.1) P-interaction by sex = 0.13 DMPs odds ratio = 1.7 (0.5, 5.5) DEPs odds ratio = 3.0 (1.1, 8.2)	Psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and after additional adjustment for birth weight, gestational age, and blood lead levels	Composite ADHD indicator is based on at least two of the following: Child Behavior Checklist ADHD scale = borderline range, Conners' Kiddie Continuous Performance Test ADHD Confidence Index ≥ 60%, and Hillside ADHD scale ≥ 75%
Marks et al. (2010)	"	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	25/316 (7.9%)	DAPs odds ratio = 1.0 (0.5, 2.0) DMPs odds ratio = 0.8 (0.4, 1.5) DEPs odds ratio = 2.0 (1.1, 3.6)	Maternal total DAPs, psychometrician, age at assessment, sex, child care, breast-feeding, maternal education, maternal depressive symptoms, and maternal Peabody Picture Vocabulary Test score Results were similar with creatinine-adjusted DAPs and without adjustment for maternal DAPs	–

Bouchard et al. (2011)	Wechsler working memory at 7 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	267 first half 279 second half 298 averaged	DAPs beta, first half of pregnancy = –1.6 (–4.2, 1.0) DAPs beta, second half of pregnancy = –3.0 (–6.4, 0.4) DAPs beta, pregnancy average = –4.3 (–7.7, –0.9) DMPs beta, pregnancy average = –4.0 (–7.1, –1.0) DEPs beta, pregnancy average = –0.4 (–3.5, 2.7)	Infant-Toddler Home Observation for Measurement of the Environment score at 6 months, maternal education, and maternal intelligence No difference after additional adjustment for maternal levels of polybrominated diphenyl ethers, polychlorinated biphenyls, dichlorodiphenyltrichloroethane/ dichlorodiphenyldichloroethylene, and lead, prenatal DAPs (in analyses of child DAPs), use of creatinine-adjusted DAPs, stratification by sex, or restriction to children tested in Spanish	Working memory = Digit Span and Letter-Number Sequencing subtests; scores standardized against U.S. population-based norms for English- and Spanish-speaking children Estimates of association did not differ significantly ($P = 0.10$) between prenatal and postnatal DAP concentrations
Bouchard et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	265 at 6 months 274 at 12 months 274 at 24 months 231 at 42 months 273 at 60 months 245 at all ages	Beta = –1.7 (–3.9, 0.5) Beta = 0.9 (–1.4, 3.2) Beta = –0.4 (–2.7, 1.9) Beta = 0.8 (–1.7, 3.3) Beta = 2.0 (–0.1, 4.0) Beta for area under curve = 1.6 (–2.2, 5.4)	"	Area under the curve = cumulative DAP level between 6 and 60 months
Bouchard et al. (2011)	Wechsler processing speed at 7 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	268 first half 280 second half 298 averaged	DAPs beta, first half of pregnancy = –1.5 (–3.9, 0.9) DAPs beta, second half of pregnancy = –2.6 (–5.9, 0.7) DAPs beta, pregnancy average = –3.4 (–6.8, –0.1) DMPs beta, pregnancy average = –1.8 (–4.8, 1.2) DEPs beta, pregnancy average = –4.0 (–7.0, –1.0)	"	Processing speed = Coding and Symbol Search subtests; scores standardized against U.S. population-based norms for English- and Spanish-speaking children Estimates of association did not differ significantly ($P = 0.24$) between prenatal and postnatal DAP concentrations; interaction term between mean prenatal DAP level and AUC was not statistically significant ($P > 0.15$)
Bouchard et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	266 at 6 months 274 at 12 months 274 at 24 months 231 at 42 months 273 at 60 months 246 at all ages	Beta = –0.3 (–2.5, 1.8) Beta = 1.6 (–0.6, 3.8) Beta = –2.0 (–4.3, 0.2) Beta = –1.1 (–3.6, 1.3) Beta = 0.7 (–1.3, 2.7) Beta for area under curve = –1.3 (–4.9, 2.3)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Bouchard et al. (2011)	Wechsler verbal comprehension at 7 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, log_{10} scale)	291 first half 309 second half 329 averaged	DAPs beta, first half of pregnancy = –2.6 (–5.1, –0.1) DAPs beta, second half of pregnancy = –3.1 (–6.4, 0.2) DAPs beta, pregnancy average = –5.3 (–8.6, –2.0) DMPs beta, pregnancy average = –4.8 (–7.8, –1.9) DEPs beta, pregnancy average = –2.0 (–5.0, 1.1)	Infant-Toddler Home Observation for Measurement of the Environment score, maternal education, maternal intelligence, and language of assessment No difference after additional adjustment for maternal levels of polybrominated diphenyl ethers, polychlorinated biphenyls, dichlorodiphenyltrichloroethane/ dichlorodiphenyldichloroethylene, and lead, prenatal DAPs (in analyses of child DAPs), use of creatinine-adjusted DAPs, stratification by sex, or restriction to children tested in Spanish	Verbal comprehension = Vocabulary and Similarities subtests; scores standardized against U.S. population-based norms for English- and Spanish-speaking children Estimates of association differed significantly ($P = 0.01$) between prenatal and postnatal DAP concentrations
Bouchard et al. (2011)	"	Child urinary DAPs (nmol/L, log_{10} scale)	294 at 6 months 303 at 12 months 303 at 24 months 259 at 42 months 302 at 60 months 271 at all ages	Beta = 0.8 (–1.4, 3.0) Beta = 2.9 (0.7, 5.2) Beta = –0.8 (–3.1, 1.5) Beta = 0.2 (–2.2, 2.6) Beta = 0.4 (–1.6, 2.5) Beta for area under curve = 0.8 (–3.0, 4.6)	"	–
Bouchard et al. (2011)	Wechsler perceptual reasoning at 7 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, log_{10} scale)	292 first half 309 second half 329 averaged	DAPs beta, first half of pregnancy = –1.2 (–4.1, 1.7) DAPs beta, second half of pregnancy = –2.4 (–6.3, 1.4) DAPs beta, pregnancy average = –4.0 (–7.9, –0.1) DMPs beta, pregnancy average = –3.3 (–6.7, 0.2) DEPs beta, pregnancy average = –2.1 (–5.6, 1.5)	Infant-Toddler Home Observation for Measurement of the Environment score, maternal education, and maternal intelligence No difference after additional adjustment for maternal levels of polybrominated diphenyl ethers, polychlorinated biphenyls, dichlorodiphenyltrichloroethane/ dichlorodiphenyldichloroethylene, and lead, prenatal DAPs (in analyses of child DAPs), use of creatinine-adjusted DAPs, stratification by sex, or restriction to children tested in Spanish	Perceptual reasoning = Block Design and Matrix Reasoning subtests; scores standardized against U.S. population-based norms for English- and Spanish-speaking children Estimates of association did not differ significantly ($P = 0.19$) between prenatal and postnatal DAP concentrations
Bouchard et al. (2011)	"	Child urinary DAPs (nmol/L, log_{10} scale)	294 at 6 months 303 at 12 months 303 at 24 months 259 at 42 months 302 at 60 months 271 at all ages	Beta = –2.4 (–4.9, 0.1) Beta = 1.9 (–0.8, 4.5) Beta = –0.7 (–3.4, 2.0) Beta = –0.3 (–3.0, 2.5) Beta = 2.3 (–0.1, 4.7) Beta for area under curve = 0.5 (–3.8, 4.8)	"	–

Bouchard et al. (2011)	Wechsler full-scale intelligence quotient at 7 years	Maternal prenatal urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	266 first half 279 second half 297 averaged	DAPs beta, first half of pregnancy = –2.4 (–4.9, 0.2) DAPs beta, second half of pregnancy = –3.5 (–6.9, –0.1) DAPs beta, pregnancy average = –5.6 (–9.0, –2.2) DMPs beta, pregnancy average = –4.7 (–7.7, –1.6) DEPs beta, pregnancy average = –2.8 (–5.6, 0.3)	Infant-Toddler Home Observation for Measurement of the Environment score, maternal education, maternal intelligence, and language of assessment No difference after additional adjustment for maternal levels of polybrominated diphenyl ethers, polychlorinated biphenyls, dichlorodiphenyltrichloroethane/ dichlorodiphenyldichloroethylene, and lead, prenatal DAPs (in analyses of child DAPs), use of creatinine-adjusted DAPs, stratification by sex, or restriction to children tested in Spanish	Estimates of association differed significantly ($P = 0.03$) between prenatal and postnatal DAP concentrations
Bouchard et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	265 at 6 months 273 at 12 months 273 at 24 months 231 at 42 months 272 at 60 months 245 at all ages	Beta = –0.9 (–3.2, 1.3) Beta = 2.7 (0.3, 5.1) Beta = –1.5 (–3.9, 0.9) Beta = 0.2 (–2.4, 2.8) Beta = 1.7 (–0.4, 3.9) Beta for area under curve = 0.6 (–3.2, 4.4)	"	–
Quiros-Alcala et al. (2011)	Respiratory sinus arrhythmia, resting (index) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.05 (–0.33, 0.24) Beta = –0.11 (–0.43, 0.21) Beta = 0.19 (–0.35, 0.73) Beta = 0.14 (–0.22, 0.49)	Sex, exact age at assessment, breast-feeding duration, location of assessment, psychometrician, and both prenatal and child DAPs Results based on creatinine-adjusted metabolite levels were "similar … although some associations were attenuated" (results NR)	Resting conditions at 6 months and 1 year: listening to digitally recorded lullabies Resting conditions at 3.5 and 5 years: listening to a story read aloud For resting measures, the "only significant association in both the unadjusted and creatinine-adjusted models was for child [DEP] concentrations and high [pre-ejection period] resting measures (less sympathetic activation) in 1-year-olds."
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.07 (–0.34, 0.19) Beta = –0.10 (–0.39, 0.20) Beta = 0.13 (–0.39, 0.64) Beta = 0.02 (–0.30, 0.34)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.06 (–0.37, 0.25) Beta = –0.13 (–0.47, 0.21) Beta = 0.22 (–0.30, 0.74) Beta = 0.17 (–0.17, 0.51)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.27 (–0.48, –0.06) Beta = –0.06 (–0.28, 0.16) Beta = –0.13 (–0.46, 0.20) Beta = –0.11 (–0.34, 0.12)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.24 (–0.42, –0.05) Beta = –0.06 (–0.26, 0.13) Beta = –0.15 (–0.46, 0.17) Beta = –0.13 (–0.34, 0.09)	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 149 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.13 (–0.34, 0.09) Beta = –0.03 (–0.29, 0.24) Beta = –0.05 (–0.37, 0.27) Beta = 0.03 (–0.17, 0.23)	"	–
Quiros-Alcala et al. (2011)	Heart rate, resting (beats per minute) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = –1.09 (–4.96, 2.78) Beta = –2.50 (–6.72, 1.73) Beta = –3.82 (–8.74, 1.10) Beta = 1.36 (–1.89, 4.60)	"	No significant associations were found between cumulative measures of prenatal or childhood metabolite levels (based on area under the concentration-time curve calculations) and resting or reactive measures at age 5 years, except between creatinine-unadjusted cumulative prenatal DEP levels and resting heart rate (beta = –3.19 [–6.29, –0.09])
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = –1.19 (–4.71, 2.33) Beta = –2.48 (–6.39, 1.43) Beta = –3.11 (–7.81, 1.60) Beta = 1.96 (–0.93, 4.85)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = –0.38 (–4.51, 3.75) Beta = –1.63 (–6.12, 2.86) Beta = –3.78 (–8.51, 0.95) Beta = –0.77 (–3.87, 2.33)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 1.39 (–1.43, 4.21) Beta = 0.38 (–2.58, 3.34) Beta = 2.17 (–0.83, 5.16) Beta = –1.14 (–3.21, 0.93)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 1.41 (–1.11, 3.94) Beta = 0.44 (–2.14, 3.02) Beta = 2.28 (–0.58, 5.14) Beta = –1.13 (–3.09, 0.84)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	143 at 6 months 147 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 0.62 (–2.23, 3.47) Beta = –0.02 (–3.57, 3.53) Beta = 1.58 (–1.38, 4.53) Beta = –0.14 (–1.92, 1.65)	"	–
Quiros-Alcala et al. (2011)	Pre-ejection period, resting (milliseconds) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.67 (–4.11, 2.76) Beta = 3.72 (–0.09, 7.53) Beta = 1.27 (–1.92, 4.47) Beta = –0.86 (–3.11, 1.39)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.77 (–3.90, 2.35) Beta = 3.77 (0.21, 7.33) Beta = 1.04 (–2.01, 4.09) Beta = –1.18 (–3.18, 0.83)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = 1.25 (–2.46, 4.96) Beta = 2.74 (–1.20, 6.68) Beta = 1.03 (–2.01, 4.07) Beta = 0.39 (–1.75, 2.53)	"	–

Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.31 (–2.83, 2.22) Beta = 1.27 (–1.39, 3.92) Beta = 0.57 (–1.38, 2.51) Beta = 0.35 (–1.09, 1.79)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.06 (–2.27, 2.16) Beta = 0.34 (–1.99, 2.68) Beta = 0.74 (–1.11, 2.59) Beta = 0.27 (–1.10, 1.63)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	136 at 6 months 141 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.59 (–3.15, 1.98) Beta = 4.33 (1.24, 7.42) Beta = –0.96 (–2.87, 0.94) Beta = 0.70 (–0.53, 1.93)	"	–
Quiros-Alcala et al. (2011)	Respiratory sinus arrhythmia, reactive (index) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.17 (–0.36, 0.03) Beta = 0.24 (0.03, 0.46) Beta = 0.06 (–0.23, 0.34) Beta = –0.08 (–0.25, 0.08)	Sex, exact age at assessment, breast-feeding duration, location of assessment, psychometrician, and both prenatal and child DAPs "Associations between reactivity scores and creatinine-adjusted prenatal [DMP] and DAP levels were similar" (results NR)	Challenging conditions at 6 months and 1 year: watching a jack-in-the-box wound up and jumping out of the box (social/startle), listening to a digitally recorded sick baby crying (emotion), and feeling a vibrator on the leg (physical) Challenging conditions at 3.5 and 5 years: answering questions (social), watching a scary video clip (emotion), tasting concentrated lemon juice on the tongue (physical), and repeating a series of numbers (cognitive)
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.15 (–0.33, 0.03) Beta = 0.25 (0.05, 0.45) Beta = 0.07 (–0.21, 0.34) Beta = –0.04 (–0.18, 0.11)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.09 (–0.31, 0.12) Beta = 0.01 (–0.22, 0.24) Beta = 0.02 (–0.25, 0.29) Beta = –0.01 (–0.17, 0.14)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.10 (–0.24, 0.05) Beta = 0.01 (–0.14, 0.16) Beta = –0.03 (–0.20, 0.14) Beta = 0.06 (–0.04, 0.16)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = –0.08 (–0.21, 0.05) Beta = –0.01 (–0.14, 0.12) Beta = –0.02 (–0.19, 0.14) Beta = 0.06 (–0.04, 0.16)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	141 at 6 months 147 at 1 year 95 at 3.5 years 270 at 5 years	Beta = 0.00 (–0.14, 0.15) Beta = 0.09 (–0.09, 0.27) Beta = 0.01 (–0.16, 0.18) Beta = –0.02 (–0.11, 0.07)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Quiros-Alcala et al. (2011)	Heart rate, reactive (beats per minute) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 0.62 (–1.37, 2.62) Beta = –0.20 (–2.38, 1.98) Beta = –0.51 (–2.31, 1.28) Beta = 0.42 (–0.87, 1.72)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 0.52 (–1.31, 2.36) Beta = –0.32 (–2.34, 1.70) Beta = –0.44 (–2.16, 1.27) Beta = 0.19 (–0.96, 1.35)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 0.44 (–1.69, 2.57) Beta = 0.13 (–2.16, 2.42) Beta = –0.69 (–2.41, 1.02) Beta = 0.22 (–1.01, 1.45)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 1.20 (–0.26, 2.67) Beta = –0.36 (–1.88, 1.15) Beta = 0.08 (–1.01, 1.17) Beta = –0.09 (–0.91, 0.73)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 0.78 (–0.54, 2.09) Beta = –0.13 (–1.46, 1.19) Beta = –0.02 (–1.07, 1.02) Beta = 0.08 (–0.71, 0.86)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	142 at 6 months 145 at 1 year 95 at 3.5 years 271 at 5 years	Beta = 1.23 (–0.24, 2.71) Beta = –1.09 (–2.89, 0.71) Beta = 0.00 (–1.07, 1.08) Beta = –0.30 (–1.00, 0.41)	"	–
Quiros-Alcala et al. (2011)	Pre-ejection period, reactive (milliseconds) at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = 1.23 (–0.07, 2.54) Beta = –1.07 (–2.56, 0.41) Beta = 0.27 (–0.67, 1.21) Beta = –0.35 (–0.85, 0.16)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DMPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = 1.21 (0.03, 2.40) Beta = –1.00 (–2.39, 0.38) Beta = 0.23 (–0.67, 1.13) Beta = –0.32 (–0.77, 0.14)	"	–
Quiros-Alcala et al. (2011)	"	Maternal prenatal urinary DEPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = 0.07 (–1.37, 1.51) Beta = –0.08 (–1.66, 1.49) Beta = 0.18 (–0.72, 1.07) Beta = –0.26 (–0.74, 0.22)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.03 (–1.00, 0.93) Beta = 0.64 (–0.40, 1.67) Beta = –0.23 (–0.80, 0.34) Beta = 0.18 (–0.14, 0.50)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DMPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = 0.05 (–0.80, 0.89) Beta = 0.37 (–0.53, 1.28) Beta = –0.20 (–0.74, 0.35) Beta = 0.15 (–0.16, 0.46)	"	–
Quiros-Alcala et al. (2011)	"	Child urinary DEPs (nmol/L, $\log_{10}$ scale)	135 at 6 months 137 at 1 year 94 at 3.5 years 269 at 5 years	Beta = –0.11 (–1.10, 0.88) Beta = 0.75 (–0.50, 2.00) Beta = –0.21 (–0.77, 0.36) Beta = 0.23 (–0.05, 0.50)	"	–

Quiros-Alcala et al. (2011)	Autonomic nervous system profile at 6 months and 1, 3.5, and 5 years	Maternal prenatal urinary DAPs (nmol/L)	6 months: 22 coactivation 43 coinhibition 41 reciprocal parasympathetic activation 20 reciprocal sympathetic activation 1 year: 35 coactivation 33 coinhibition 43 reciprocal parasympathetic activation 21 reciprocal sympathetic activation 3.5 years: 11 coactivation 26 coinhibition 14 reciprocal parasympathetic activation 40 reciprocal sympathetic activation 5 years: 47 coactivation 75 coinhibition 41 reciprocal parasympathetic activation 99 reciprocal sympathetic activation	 Geometric mean = 198.3 (143.6, 273.8) Geometric mean = 110.0 (69.9, 173.1) Geometric mean = 160.8 (113.8, 227.3) Geometric mean = 110.0 (69.9, 173.1) F = 1.53, $P = 0.21$ Geometric mean = 216.4 (157.0, 298.4) Geometric mean = 141.8 (100.8, 199.5) Geometric mean = 173.4 (125.2, 240.3) Geometric mean = 143.1 (77.4, 264.6) F = 1.12, $P = 0.34$ Geometric mean = 198.0 (101.9, 384.7) Geometric mean = 185.9 (117.0, 295.4) Geometric mean = 128.3 (72.4, 227.5) Geometric mean = 121.2 (96.1, 152.9) F = 1.58, $P = 0.20$ Geometric mean = 138.0 (102.0, 186.6) Geometric mean = 132.2 (106.3, 164.3) Geometric mean = 115.0 (85.8, 154.1) Geometric mean = 149.7 (126.4, 177.3) F = 0.83, $P = 0.48$	None "Results were similar when using creatinine-adjusted prenatal concentrations" (NR)	Coactivation profile: activation of both sympathetic and parasympathetic nervous systems during challenge tasks compared with rest Coinhibition profile: inhibition of both sympathetic and parasympathetic nervous systems during challenge tasks compared with rest Reciprocal parasympathetic nervous system activation and sympathetic nervous system withdrawal Reciprocal sympathetic nervous system activation and parasympathetic nervous system withdrawal Frequencies of coactivation and coinhibition at 6 months are taken from Table 4b of manuscript (inconsistent with Table 4a)

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Quiros-Alcala et al. (2011)	"	Child urinary DAPs (nmol/L)	6 months: 22 coactivation 43 coinhibition 41 reciprocal parasympathetic activation 20 reciprocal sympathetic activation 1 year: 35 coactivation 33 coinhibition 43 reciprocal parasympathetic activation 21 reciprocal sympathetic activation 3.5 years: 11 coactivation 26 coinhibition 14 reciprocal parasympathetic activation 40 reciprocal sympathetic activation 5 years: 47 coactivation 75 coinhibition 41 reciprocal parasympathetic activation 99 reciprocal sympathetic activation	 Geometric mean = 46.3 (28.2, 76.2) Geometric mean = 58.8 (39.5, 87.3) Geometric mean = 38.8 (24.9, 60.5) Geometric mean = 88.4 (40.3, 193.9) F = 1.79, $P = 0.15$ Geometric mean = 34.5 (20.4, 58.2) Geometric mean = 58.0 (32.2, 104.5) Geometric mean = 68.6 (42.2, 111.8) Geometric mean = 46.9 (21.0, 105.0) F = 1.25, $P = 0.29$ Geometric mean = 93.3 (23.2, 376.0) Geometric mean = 77.2 (45.2, 131.8) Geometric mean = 73.6 (37.1, 146.3) Geometric mean = 132.1 (78.9, 221.3) F = 0.87, $P = 0.46$ Geometric mean = 99.4 (66.2, 149.4) Geometric mean = 110.0 (79.2, 153.0) Geometric mean = 124.8 (73.7, 211.4) Geometric mean = 90.4 (68.7, 119.0) F = 0.57, $P = 0.63$	None F = 3.18, $P = 0.03$ for creatinine-adjusted child urinary DAPs and reciprocal sympathetic activation with parasympathetic withdrawal at 6 months	No significant differences were observed in autonomic nervous system profiles between children with consistently high (top 10%) vs. consistently low (bottom 10%) prenatal and/or childhood DAP levels
Lizardi et al. (2008)	Trail Making Test B (seconds) at ~ 7 years	Child urinary DAPs ≥ 25 vs. < 25 µg/L in original screening sample	24 detectable (≥ 25 µg/L) 22 non-detectable (< 25 µg/L)	Mean = 283 (224, 341) Mean = 204 (172, 236) $P = 0.01$	None	One child in each exposure group with a significantly higher urinary DAP level (519 µg/L and 850 µg/L) was excluded from analysis

Lizardi et al. (2008)	Wechsler Intelligence Scale for Children–Third Edition Short Form, Children's Memory Scale, Wisconsin Card Sorting Test, Trail Making Test A, Child Behavior Checklist/4–18, and Teacher Report Form at ~ 7 years	"	"	"No significant effects" (results NR)	"	One child in each exposure group with a significantly higher urinary DAP level (519 µg/L and 850 µg/L) was excluded from analysis
Lizardi et al. (2008)	Wisconsin Card Sorting Test measures at ~ 7 years	Child urinary DAPs (µg/L) in contemporaneous sample	48 46 after exclusion of outliers	Number of errors made: correlation = 0.31, $P = 0.03$ Number of perseverative responses: correlation = 0.34, $P = 0.01$ Number of perseverative errors: correlation = 0.35, $P = 0.01$ Conceptual level responses provided: correlation = 0.38, $P = 0.01$ Failure to maintain set: correlation = 0.38, $P = 0.02$ After exclusion of one child in each exposure group with a significantly high urinary DAP level: "no significant correlations" (results NR)	"	–
Lizardi et al. (2008)	Wechsler Intelligence Scale for Children–Third Edition Short Form, Children's Memory Scale, Trail Making Test A and B, Child Behavior Checklist/4–18, and Teacher Report Form at ~ 7 years	"	"	"No significant correlations ($p < .05$)" (results NR)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Bouchard et al. (2010)	ADHD by diagnostic criteria at 8–15 years	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	119/1139 (10.4%)	DAPs odds ratio = 1.21 (0.97, 1.51) DMPs odds ratio = 1.55 (1.14, 2.10) DEPs odds ratio = 0.94 (0.69, 1.28)	Gender, age, race/ethnicity, ratio of family income to poverty level, fasting duration, and logarithmically transformed urinary creatinine concentration No change after additional adjustment for year of data collection, blood lead concentration, maternal age at birth, or maternal smoking pregnancy, or after exclusion of children taking ADHD medication	Diagnosis of ADHD is based on the presence during previous 12 months of symptoms related to inattention, hyperactivity, and impulsivity, with significant impairment in ≥ settings (e.g., at school and at home); no requirement that symptoms occur without another neuropsychiatric disorder or that symptoms were present before 7 years of age Results were similar when using creatinine-adjusted DAP, DMP, and DEP concentrations (results NR)
Bouchard et al. (2010)	"	Child urinary dimethylthio-phosphate (nmol/g creatinine)	407 below detection limit 366 < median (30.4 nmol/g creatinine) 366 ≥ median (30.4 nmol/g creatinine)	Odds ratio = referent Odds ratio = 1.05 (0.57, 1.95) Odds ratio = 1.93 (1.23, 3.02)	Gender, age, race/ethnicity, ratio of family income to poverty level, and fasting duration No change after additional adjustment for year of data collection, blood lead concentration, maternal age at birth, or maternal smoking during pregnancy, or after exclusion of children taking ADHD medication	–
Bouchard et al. (2010)	ADHD by diagnostic criteria or medication use at 8–15 years	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	148/1139 (13.0%)	DAPs odds ratio = 1.35 (1.10, 1.67) DMPs odds ratio = 1.72 (1.31, 2.28) DEPs odds ratio = 0.80 (0.60, 1.05)	Gender, age, race/ethnicity, ratio of family income to poverty level, fasting duration, and logarithmically transformed urinary creatinine concentration No change after additional adjustment for year of data collection, blood lead concentration, maternal age at birth, or maternal smoking during pregnancy	–
Bouchard et al. (2010)	"	Child urinary dimethylthio-phosphate (nmol/g creatinine)	407 below detection limit 366 < median (30.4 nmol/g creatinine) 366 ≥ median (30.4 nmol/g creatinine)	Odds ratio = referent Odds ratio = 1.22 (0.65, 2.27) Odds ratio = 2.12 (1.32, 3.41)	Gender, age, race/ethnicity, ratio of family income to poverty level, and fasting duration No change after additional adjustment for year of data collection, blood lead concentration, maternal age at birth, or maternal smoking during pregnancy	–

Bouchard et al. (2010)	Hyperactive/ impulsive ADHD subtype at 8–15 years	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	21/1139 (1.8%)	DAPs odds ratio = 1.85 (1.04, 3.27) DMPs odds ratio = 2.13 (1.08, 4.20) DEPs odds ratio = 2.15 (1.06, 4.40)	Gender, age, race/ethnicity, ratio of family income to poverty level, fasting duration, and logarithmically transformed urinary creatinine concentration No change after additional adjustment for year of data collection, blood lead concentration, maternal age at birth, or maternal smoking pregnancy, or after exclusion of children taking ADHD medication	–
Bouchard et al. (2010)	Inattentive ADHD subtype at 8–15 years	"	69/1139 (6.1%)	DAPs odds ratio = 1.14 (0.81, 1.61) DMPs odds ratio = 1.47 (0.99, 2.19) DEPs odds ratio = 0.70 (0.49, 1.01)	"	–
Bouchard et al. (2010)	Combined ADHD subtype at 8–15 years	"	29/1139 (2.5%)	DAPs odds ratio = 1.05 (0.51, 2.16) DMPs odds ratio = 1.30 (0.48, 3.48) DEPs odds ratio = 1.22 (0.59, 2.50)	"	–
Guodong et al. (2012)	Gesell motor behavior at 23–25 months	Child urinary DAPs (nmol/g creatinine, $\log_{10}$ scale)	301 300 normal, 1 (0.3%) with developmental delay	DAPs beta = 0.30 (− 1.40, 1.99) DMPs beta = − 1.25 (− 2.98, 0.47) DEPs beta = 0.32 (− 1.37, 2.01)	Child sex, maternal education level, and household income	Motor behavior includes locomotion, reaching, balance, comprehension, drawing, and hand control Standardized to mean ± SD of 100 ± 15, with < 85 indicating developmental delay
Guodong et al. (2012)	Gesell adaptive behavior at 23–25 months	"	301 297 normal, 4 (1.3%) with developmental delay	DAPs beta = 1.71 (− 1.15, 4.57) DMPs beta = 2.53 (− 0.05, 5.10) DEPs beta = − 0.41 (− 3.22, 2.39)	"	Adaptive behavior includes hand-eye coordination, imitation, object recovery, comprehension, discriminative performance, perception, completion, and number conception Standardized to mean ± SD of 100 ± 15, with < 85 indicating developmental delay
Guodong et al. (2012)	Gesell language behavior at 23–25 months	"	301 282 normal, 19 (6.3%) with developmental delay	DAPs beta = 2.79 (− 1.01, 6.60) DMPs beta = 2.83 (− 0.60, 6.26) DEPs beta = − 0.29 (− 4.02, 3.44)	"	Language behavior includes vocabulary, word comprehension, conversation, and word production Standardized to mean ± SD of 100 ± 15, with < 85 indicating developmental delay
Guodong et al. (2012)	Gesell social behavior at 23–25 months	"	301 293 normal, 8 (2.7%) with developmental delay	DAPs beta = − 0.66 (− 2.12, 0.79) DMPs beta = − 0.48 (− 1.93, 0.97) DEPs beta = − 0.93 (− 2.40, 0.54)	"	Personal and social behavior includes reactions to people, personal habits, initiative and independence, play responses, and acquired information Standardized to mean ± SD of 100 ± 15, with < 85 indicating developmental delay

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Yolton et al. (2013)	NICU Network attention subscale at 5 weeks	Maternal 16- and 26-week average prenatal urinary DEPs (nmol/g creatinine, $\log_2$ scale)	350	Beta = 0.066, SE = 0.033, $P < 0.05$	Infant age at exam and race	Results are presented only for statistically significant associations Positive coefficient = increased attention
Yolton et al. (2013)	NICU Network lethargy subscale at 5 weeks	Maternal 16-week prenatal urinary DEPs (nmol/g creatinine, $\log_2$ scale)	"	Beta = –0.069, SE = 0.034, $P = 0.04$	Infant age at exam, race, birth weight, and maternal consumption of fresh fruits and vegetables	Negative coefficient = decreased lethargy
Yolton et al. (2013)	NICU Network hypotonia subscale at 5 weeks	"	"	Beta = –0.101, SE = 0.045, $P = 0.03$	Infant age at exam, race, and maternal body mass index	Negative coefficient = decreased hypotonia
Yolton et al. (2013)	NICU Network autonomic stress subscale at 5 weeks	Maternal 26-week prenatal urinary DAPs (nmol/g creatinine, $\log_2$ scale)	"	Beta = –0.010, SE = 0.004, $P = 0.01$	Infant age at exam, race, birth weight, and blood lead level	Negative coefficient = decreased autonomic stress
Yolton et al. (2013)	All other NICU Network subscales at 5 weeks	All other maternal prenatal urinary DAPs, DMPs, and DEPs (nmol/g creatinine, $\log_2$ scale) at 16 weeks, 26 weeks, or averaged	"	Not statistically significant ($P > 0.05$)	NR	–
Yolton et al. (2013)	NICU Network profile at 5 weeks	Maternal prenatal urinary DAPs (nmol/g creatinine, $\log_2$ scale)	157 (45%) social/easy-going	Odds ratio = referent	Infant age at exam, race, maternal weight gain during pregnancy, and maternal body mass index	Profiles identified using latent profile analysis of patterns across NICU Network Neurobehavioral Scale dimensions
			83 (31%) high-arousal/difficult	Odds ratio, 16- and 26-week mean = 1.14 (0.98, 1.32) Odds ratio, 16-week = 1.02 (0.91, 1.15) Odds ratio, 26-week = 1.13 (0.99, 1.27)		
			110 (24%) hypotonic	Odds ratio, 16- and 26-week mean = 1.02 (0.87, 1.19) Odds ratio, 16-week = 0.90 (0.79, 1.03) Odds ratio, 26-week = 1.13 (0.99, 1.29)		

Yolton et al. (2013)	"	Maternal prenatal urinary DMPs (nmol/g creatinine, $\log_2$ scale)	157 (45%) social/ easy-going 83 (31%) high-arousal/difficult 110 (24%) hypotonic	Odds ratio = referent Odds ratio, 16- and 26-week mean = 1.11 (0.97, 1.26) Odds ratio, 16-week = 1.00 (0.90, 1.10) Odds ratio, 26-week = 1.12 (1.00, 1.25) Odds ratio, 16- and 26-week mean = 0.99 (0.86, 1.13) Odds ratio, 16-week = 0.90 (0.80, 1.00) Odds ratio, 26-week = 1.12 (0.99, 1.26)	"	–
Yolton et al. (2013)	"	Maternal prenatal urinary DEPs (nmol/g creatinine, $\log_2$ scale)	157 (45%) social/ easy-going 83 (31%) high-arousal/difficult 110 (24%) hypotonic	Odds ratio = referent Odds ratio, 16- and 26-week mean = 1.03 (0.92, 1.15) Odds ratio, 16-week = 0.98 (0.89, 1.08) Odds ratio, 26-week = 1.03 (0.95, 1.12) Odds ratio, 16- and 26-week mean = 0.96 (0.86, 1.09) Odds ratio, 16-week = 0.89 (0.81, 0.99) Odds ratio, 26-week = 1.03 (0.94, 1.13)	"	–
Oulhote and Bouchard (2013)	Strengths and Difficulties Questionnaire total difficulties score ≥ 17 at ages 6–11 years	Child urinary DAPs, DMPs, or DEPs (nmol/L, $\log_{10}$ scale)	69 (6.8%) overall 48 boys 21 girls 779 subjects in analysis	DAPs odds ratio, total = 0.6 (0.3, 1.3) DAPs odds ratio, boys = 0.5 (0.2, 1.8) DAPs odds ratio, girls = 0.6 (0.3, 1.6) P-interaction by sex = 0.97 DMPs odds ratio = 0.8 (0.4, 1.6) DEPs odds ratio = 0.3 (0.1, 1.8)	Sex, age, race/ethnicity, income, parental education, maternal smoking during pregnancy, birth weight, blood lead levels, urinary creatinine, body mass index, and fasting status No change when using creatinine-standardized metabolite concentrations, not adjusting for blood lead, or not weighting to account for survey design	Questionnaire is designed for screening of mental and behavioral difficulties and strengths in population surveys; each dimension scale is scored on a scale of 0–0, and total difficulties are calculated by summing four dimension scales (total of 0–40) Prosocial behavior not analyzed because "too few children had high scores"
Oulhote and Bouchard (2013)	Strengths and Difficulties Questionnaire conduct problems score ≥ 4 at ages 6–11 years	Child urinary DAPs (nmol/L, $\log_{10}$ scale)	78 (8.0%) overall 53 boys 25 girls 779 subjects in analysis	DAPs odds ratio, total = 0.6 (0.3, 1.3) DAPs odds ratio, boys = 0.7 (0.3, 1.7) DAPs odds ratio, girls = 0.4 (0.1, 2.0) P-interaction by sex = 0.58	"	–
Oulhote and Bouchard (2013)	Strengths and Difficulties Questionnaire emotional symptoms score ≥ 5 at ages 6–11 years	"	97 (9.1%) overall 47 boys 50 girls 779 subjects in analysis	DAPs odds ratio, total = 1.1 (0.5, 2.2) DAPs odds ratio, boys = 1.0 (0.4, 3.0) DAPs odds ratio, girls = 1.0 (0.5, 2.0) P-interaction by sex = 0.96	"	–

(Continued)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Oulhote and Bouchard (2013)	Strengths and Difficulties Questionnaire hyperactivity/ inattention score ≥ 7 at ages 6–11 years	"	109 (11.1%) overall 76 boys 33 girls 779 subjects in analysis	DAPs odds ratio, total = 0.8 (0.3, 2.0) DAPs odds ratio, boys = 0.9 (0.4, 2.4) DAPs odds ratio, girls = 0.4 (0.1, 1.8) P-interaction by sex = 0.21	"	–
Oulhote and Bouchard (2013)	Strengths and Difficulties Questionnaire peer problems score ≥ 4 at ages 6–11 years	"	71 (7.3%) overall 43 boys 28 girls 779 subjects in analysis	DAPs odds ratio, total = 0.8 (0.3, 2.0) DAPs odds ratio, boys = 0.8 (0.3, 2.3) DAPs odds ratio, girls = 0.6 (0.2, 2.7) P-interaction by sex = 0.75	"	–
Fortenberry et al. (2014)	Conners parent-rated ADHD index at 6–11 years	Maternal prenatal TCPy (ng/mL), tertiles 2 and 3 vs. 1	187 total 80 males 97 females	Tertile 2 beta, total = 2.61 (− 1.54, 6.75) Tertile 3 beta, total = 4.00 (− 0.91, 8.90) P-trend, total = 0.11 Tertile 2 beta, males = 2.32 (− 2.55, 7.20) Tertile 3 beta, males = 5.55 (− 0.19, 11.3) P-trend, males = 0.06 Tertile 2 beta, females = 1.63 (− 5.55, 8.82) Tertile 3 beta, females = 0.17 (− 8.28, 8.63) P-trend, females = 0.96	Child sex, maternal intelligence quotient, maternal education, income, child age at testing, specific gravity, season, breast feeding, blood lead, delivery length, and delivery head circumference	No significant differences in geometric mean TCPy concentrations were detected between trimesters, but significant within-person variability was detected across trimesters (intraclass correlation = 0.29–0.32 for specific-gravity-corrected TCPy, 0.41 for uncorrected TCPy) Higher score on Conners' Parental Rating Scales–Revised ADHD Index indicates an elevated level of concern for risk of ADHD, with a score of 40–59 being average and < 40 displaying fewer concerns Various tests are used to "assess ADHD-related symptoms and are not designed as diagnostic tools, but rather for screening"
Fortenberry et al. (2014)	Conners parent-rated hyperactivity/ impulsivity ADHD at 6–11 years	"	"	Tertile 2 beta, total = − 0.56 (− 5.03, 3.91) Tertile 3 beta, total = − 0.51 (− 5.80, 4.78) P-trend, total = 0.84 Tertile 2 beta, males = − 0.17 (− 6.63, 6.29) Tertile 3 beta, males = 1.25 (− 6.36, 8.87) P-trend, males = 0.76 Tertile 2 beta, females = 0.33 (− 6.44, 7.10) Tertile 3 beta, females = − 3.81 (− 11.8, 4.16) P-trend, females = 0.35	"	Scale based on *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*, with scores ranging between 0 and 9 and scores ≥ 6 suggesting a possible diagnosis

Fortenberry et al. (2014)	Conners parent-rated inattention ADHD at 6–11 years	"	"	Tertile 2 beta, total = 2.37 (− 1.79, 6.53) Tertile 3 beta, total = 2.45 (− 2.47, 7.37) P-trend, total = 0.31 Tertile 2 beta, males = 2.33 (− 2.36, 7.02) Tertile 3 beta, males = 2.63 (− 2.89, 8.16) P-trend, males = 0.32 Tertile 2 beta, females = 1.19 (− 6.09, 8.47) Tertile 3 beta, females = − 0.07 (− 8.64, 8.50) P-trend, females = 0.99	"	Scale based on *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*, with scores ranging between 0 and 9 and scores ≥ 6 suggesting a possible diagnosis
Fortenberry et al. (2014)	Conners parent-rated combined ADHD at 6–11 years	"	"	Tertile 2 beta, total = 1.23 (− 2.89, 5.35) Tertile 3 beta, total = 1.10 (− 3.77, 5.98) P-trend, total = 0.64 Tertile 2 beta, males = 0.80 (− 4.48, 6.09) Tertile 3 beta, males = 2.06 (− 4.17, 8.29) P-trend, males = 0.51 Tertile 2 beta, females = 1.64 (− 5.17, 8.45) Tertile 3 beta, females = − 1.83 (− 9.84, 6.19) P-trend, females = 0.66	"	Scale based on *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*, with scores ranging between 0 and 9 and scores ≥ 6 suggesting a possible diagnosis
Fortenberry et al. (2014)	Conners parent-rated global restlessness/ impulsivity index at 6–11 years	"	"	Tertile 2 beta, total = − 0.15 (− 4.57, 4.27) Tertile 3 beta, total = 0.38 (− 4.85, 5.61) P-trend, total = 0.89 Tertile 2 beta, males = 0.49 (− 5.71, 6.68) Tertile 3 beta, males = 3.78 (− 3.52, 11.1) P-trend, males = 0.32 Tertile 2 beta, females = − 0.48 (− 7.10, 6.14) Tertile 3 beta, females = − 4.90 (− 12.7, 2.89) P-trend, females = 0.22	"	Higher score on Conners' Parental Rating Scales–Revised Global Restlessness/Impulsivity Index indicates an elevated level of concern for tendencies toward hyperactivity and inattention, with a score of 40–59 being average and < 40 displaying fewer concerns

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Fortenberry et al. (2014)	Behavioral Assessment System for Children attention problems at 6–11 years	"	"	Tertile 2 beta, total = 1.79 (− 2.66, 6.24) Tertile 3 beta, total = 3.46 (− 1.81, 8.73) P-trend, total = 0.19 Tertile 2 beta, males = − 0.37 (− 7.02, 6.27) Tertile 3 beta, males = 5.59 (− 2.24, 13.4) P-trend, males = 0.18 Tertile 2 beta, females = 5.81 (− 0.75, 12.4) Tertile 3 beta, females = 1.82 (− 5.91, 9.55) P-trend, females = 0.62	"	Higher score on Behavioral Assessment for Children–Parental Rating Scales indicates elevated level of concern, with scores ≥ 59 indicating increased levels of attention/hyperactivity problems
Fortenberry et al. (2014)	Behavioral Assessment System for Children hyperactivity at 6–11 years	"	"	Tertile 2 beta, total = − 3.69 (− 7.88, 0.50) Tertile 3 beta, total = − 3.35 (− 8.31, 1.60) P-trend, total = 0.17 Tertile 2 beta, males = − 5.00 (− 12.0, 2.00) Tertile 3 beta, males = − 3.49 (− 11.7, 4.73) P-trend, males = 0.36 Tertile 2 beta, females = − 0.005 (− 5.17, 5.16) Tertile 3 beta, females = − 2.77 (− 8.84, 3.31) P-trend, females = 0.37	"	Higher score on Behavioral Assessment for Children–Parental Rating Scales indicates elevated level of concern, with scores ≥ 59 indicating increased levels of attention/hyperactivity problems
Fortenberry et al. (2014)	Conners clinical ADHD index at 6–11 years	"	"	Tertile 2 beta, total = − 3.97 (− 12.5, 4.51) Tertile 3 beta, total = 2.19 (− 8.11, 12.5) P-trend, total = 0.73 Tertile 2 beta, males = − 4.29 (− 15.8, 7.18) Tertile 3 beta, males = 0.84 (− 12.8, 14.5) P-trend, males = 0.95 Tertile 2 beta, females = 0.42 (− 13.2, 14.0) Tertile 3 beta, females = 8.55 (− 7.83, 24.9) P-trend, females = 0.31	"	Conners' Continuous Performance Test clinical index measures the likelihood of an ADHD diagnosis, with a high sensitivity (83–90%) but poorer specificity (59–61%) when compared with clinical ADHD diagnosis

Fortenberry et al. (2014)	Conners hit reaction time block change at 6–11 years	"	"	Tertile 2 beta, total = −4.59 (−9.55, 0.36) Tertile 3 beta, total = −5.10 (−11.1, 0.91) P-trend, total = 0.09 Tertile 2 beta, males = −5.10 (−13.1, 2.92) Tertile 3 beta, males = −6.86 (−16.4, 2.68) P-trend, males = 0.14 Tertile 2 beta, females = −3.79 (−10.6, 2.98) Tertile 3 beta, females = −2.33 (−10.5, 5.82) P-trend, females = 0.55	"	Hit reaction time block change is the variability of reaction time for correct responses across blocks or sections of the Conners' Continuous Performance Test, and has been indicated as a measure of vigilance or sustained attention
Zhang et al. (2014)	Neonatal Behavioral Neurological Assessment summary score at 3 days	Maternal prenatal urinary DAPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –1.78 (–2.12, –1.45) Beta = –1.47 (–1.93, –1.01) Beta = –2.03 (–2.55, –1.52)	Maternal age, education, gestational age, prenatal body mass index, and cord blood lead concentration Results were similar using creatinine-adjusted DAP concentrations	Neonatal Behavioral Neurological Assessment summary score is based on 20 items, each score from 0–2, with > 37 considered as well developed, < 34 considered as abnormal, and 34–37 considered as acceptable No evidence of departure from linearity was observed in analyses by quintile of DAPs
Zhang et al. (2014)	"	Maternal prenatal urinary DMPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.96 (–1.35, –0.57) Beta = –0.93 (–1.45, –0.40) Beta = –1.22 (–1.89, –0.55)	"	–
Zhang et al. (2014)	"	Maternal prenatal urinary DEPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.88 (–1.30, –0.47) Beta = –0.61 (–1.15, –0.07) Beta = –0.98 (–1.58, –0.39)	"	–
Zhang et al. (2014)	Neonatal Behavioral Neurological Assessment behavior score at 3 days	Maternal prenatal urinary DAPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.65 (–0.85, –0.45) Beta = –0.50 (–0.76, –0.23) Beta = –0.84 (–1.15.20.52)	"	Neonatal Behavioral Neurological Assessment behavior scale includes six items, each scored from 0–2, for a maximum of 12 (higher = better) No evidence of departure from linearity was observed in analyses by quintile of DAPs
Zhang et al. (2014)	"	Maternal prenatal urinary DMPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Not significant (results NR)	"	–
Zhang et al. (2014)	"	g	249 total 138 boys 111 girls	Beta = –0.59 (–0.79, –0.40) Beta = –0.42 (–0.67, –0.17) Beta = –0.83 (–1.15, –0.53)	"	–

(*Continued*)

Table 3. (Continued)

Reference	Outcome	Exposure	Number of subjects/events	Estimate of association (95% CI)	Adjustment factors	Comments
Zhang et al. (2014)	Neonatal Behavioral Neurological Assessment passive tone score at 3 days	Maternal prenatal urinary DAPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.22 (–0.34, –0.10) Beta = –0.21 (–0.36, –0.02) Beta = –0.21 (–0.40, –0.02)	"	Neonatal Behavioral Neurological Assessment passive tone scale includes four items, each scored from 0–2, for a maximum of 8 (higher = better) No evidence of departure from linearity was observed in analyses by quintile of DAPs
Zhang et al. (2014)	"	Maternal prenatal urinary DMPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.22 (–0.33, –0.11) Beta = –0.19 (–0.35, –0.07) Beta = –0.18 (–0.35, –0.01)	"	–
Zhang et al. (2014)	"	Maternal prenatal urinary DEPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Not significant (results NR)	"	–
Zhang et al. (2014)	Neonatal Behavioral Neurological Assessment active tone score at 3 days	Maternal prenatal urinary DAPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.48 (–0.66, –0.30) Beta = –0.46 (–0.72, –0.21) Beta = –0.51 (–0.76, –0.25)	"	Neonatal Behavioral Neurological Assessment active tone scale includes four items, each scored from 0–2, for a maximum of 8 (higher = better) No evidence of departure from linearity was observed in analyses by quintile of DAPs
Zhang et al. (2014)	"	Maternal prenatal urinary DMPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.41 (–0.57, –0.29) Beta = –0.34 (–0.58, –0.11) Beta = –0.41 (–0.65, –0.18)	"	–
Zhang et al. (2014)	"	Maternal prenatal urinary DEPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Not significant (results NR)	"	–
Zhang et al. (2014)	Neonatal Behavioral Neurological Assessment primary reflexes score at 3 days	Maternal prenatal urinary DAPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.36 (–0.51, –0.21) Beta = –0.34 (–0.55, –0.13) Beta = –0.39 (–0.61, –0.17)	"	Neonatal Behavioral Neurological Assessment primary reflexes scale includes three items, each scored from 0–2, for a maximum of 6 (higher = better) No evidence of departure from linearity was observed in analyses by quintile of DAPs
Zhang et al. (2014)	"	Maternal prenatal urinary DMPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = –0.30 (–0.44.20.17) Beta = not significant (NR) Beta = –0.34 (–0.54, –0.14)	"	–
Zhang et al. (2014)	"	Maternal prenatal urinary DEPs (μg/L, $\log_{10}$ scale)	249 total 138 boys 111 girls	Beta = not significant (NR) Beta = –0.28 (–0.48, –0.09) Beta = not significant (NR)	"	–

ADHD attention deficity/hyperactivity disorder confidence interval, *DAP* dialkyl phosphate, *DEP* diethyl phosphate, *DMP* dimethyl phosphate, *IQR* interquartile range, *MDA* malathion dicarboxylic acid, *NICU* neonatal intensive care unit, *NR* not reported, *PON1* paraoxonase 1, *SD* standard deviation, *SE* standard error, *TCPy* 3,5,6-trichloro-2-pyridinol.

chlorpyrifos detected (Table 2) (Rauh et al. 2011). However, a significant inverse association was detected with the Wechsler Working Memory Scale (parsimonious model beta = −0.006, 95% CI = −0.009, −0.002; no substantial change after further adjustment). This association was not substantially confounded (change in beta < 10%) by childhood home environment at age 3 years, based on composite indices (total Home Observation for Measurement of the Environment or HOME score, Environmental Stimulation Scale, and Parental Nurturance Scale) derived from observational interview data (Table 2) (Horton et al. 2012). Additionally, no apparent interaction was observed between chlorpyrifos and the Parental Nurturance Scale. However, the association between chlorpyrifos and Wechsler Working Memory varied by child sex, with a significant inverse association detected only among boys (beta = −2.382, 95% CI = −3.88, −0.88) and not girls (beta = −0.524, 95% CI = −1.90, 0.85).

Forty children aged 5.9–11.2 years in the CCCEH cohort with low prenatal exposure to environmental tobacco smoke (based on maternal self-report and cotinine levels < 15 ng/mL in cord plasma) and polycyclic aromatic hydrocarbons (based on maternal third-trimester personal air monitoring levels below the median of 2.26 ng/m^3), including 20 children in the highest tertile of cord plasma chlorpyrifos (≥4.39 pg/g) and 20 below the highest tertile, participated in a study of brain morphology using T1-weighted high-resolution magnetic resonance imaging (Table 2) (Rauh et al. 2012). Significant differences between chlorpyrifos exposure groups that involved primarily white matter included bilateral enlargement of the superior temporal, posterior middle temporal, and inferior postcentral gyri; right-hemisphere enlargement of the supramarginal gyrus, inferior parietal lobule, and superior frontal gyrus, gyrus rectus, cuneus, and precuneus along the mesial wall; and inward deformations in the dorsal and mesial surfaces of the left superior frontal gyrus. No significant difference was found in overall brain size by chlorpyrifos level. Wechsler Full-Scale IQ at age 7 years was positively correlated with surface measures in the bilateral superior temporal, inferior frontal, inferior precentral, and inferior postcentral gyri and the left precuneus, and inversely correlated with surface measures in the right fusiform gyrus, among children with lower cord plasma chlorpyrifos levels, but not those in the higher-exposure group. Normal sex differences in the right inferior parietal lobule, superior marginal gyrus, and mesial superior frontal gyrus were reversed among children with higher chlorpyrifos levels. "Scattered reductions" in cortical thickness in dorsal and parietal and frontal cortices were also associated with higher chlorpyrifos levels.

The major strengths and limitations of the CCCEH cohort study were discussed above in the context of analyses of birth outcomes, and apply also to the analyses of neurodevelopmental outcomes, except that selection bias due to differential participation rates of mothers by childhood neurological outcomes at age 7 years is improbable, though not impossible. For example, risk factors such as a personal or family history of neurological problems might influence the decision to participate. However, selection bias due to differential dropout rates is a greater concern in studies with relatively long follow-up. Bias in analyses of parent-reported outcomes, such as those based on the Child Behavior Checklist, is also a concern, because the completeness and accuracy of reporting may have varied by lifestyle factors related to maternal prenatal chlorpyrifos exposure. Additional limitations are the dichotomization of cord plasma chlorpyrifos levels in several analyses, which precluded exposure–response analyses, and the focus on a single OP insecticide.

Taken together, the results of neurodevelopmental studies in the CCCEH cohort suggest associations between prenatal chlorpyrifos exposure and selected adverse neurodevelopmental outcomes, with some as-yet-unexplained heterogeneity by subgroups and numerous statistically null associations. For instance, an inverse association between cord plasma chlorpyrifos levels and lower scores on the Bayley Mental Development Index was detected at 36 months among African American children, but not among Dominican children and not in either group at 12 or 24 months. In the absence of *a priori* hypotheses, it is unclear why prenatal chlorpyrifos exposure might be associated with attention problems and pervasive developmental disorder but not externalizing or internalizing behavior problems as assessed by the Child Behavior Checklist, or with working memory among boys but not overall IQ, verbal comprehension, perceptual reasoning, or processing speed as assessed by the Wechsler Intelligence Scale for Children. Given the large number of outcomes tested, at least some of the observed associations are almost certainly due to chance. Again, neither this study nor any other study of neurodevelopmental outcomes described in this review adjusted for multiple comparisons. The observed associations with brain morphology are noteworthy, but multiple comparisons are again a concern, especially given the exclusive reporting of anatomic regions where associations with chlorpyrifos exposure were observed, but not those without any such associations. Overall, the results suggesting an adverse neurodevelopmental effect of prenatal chlorpyrifos exposure cannot reliably be interpreted as causal due to methodological limitations and internal inconsistency, and require independent confirmation in other study settings.

Mount Sinai Children's Environmental Cohort Study

The Mount Sinai CECS, described earlier, administered the Brazelton Neonatal Behavioral Assessment Scale to evaluate 28 behavioral items and 18 primitive reflexes, grouped into seven clusters, in 311 neonates prior to hospital discharge at or before 5 days (Engel et al. 2007). Subsequently, the Bayley Scales of Infant Development, 2nd Edition, were administered at 12 months ($n = 200$) and 24 months ($n = 276$), and the Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition, or the Wechsler Intelligence Scale for Children, 4th Edition, was administered at ages 6–9 years ($n = 169$) (Table 1) (Engel et al. 2011). The Brazelton scale evaluates 28 behavioral items and 18 primitive reflexes, which can be scored into seven clusters: habituation, orientation, motor, range of state, regulation of state, autonomic stability, and number and type of abnormal reflexes (including plantar, Babinski, ankle clonus, rooting, sucking, glabella, passive resistance of legs, passive resistance of arms, palmar, placing, standing, walking, crawling, incurvation, tonic deviation of head and eyes, nystagmus, tonic neck reflex, and Moro reflex). The Wechsler Preschool and Primary Scale of Intelligence is used to derive composite Verbal Comprehension, Perceptual

Reasoning, Processing Speed, and Full-Scale IQ scores; the Wechsler Intelligence Scale for Children was described above for the CCCEH study.

In adjusted models, maternal prenatal urinary levels of DAPs, DMPs, and DEPs (classified as linear on the $\log_{10}$ scale or into quartiles) and detectable MDA were not significantly associated with the Brazelton habituation, orientation, motor, range of state, regulation of state, or autonomic stability clusters (Table 2) (Engel et al. 2007). However, a $\log_{10}$-unit increase in DEP levels was associated with a significantly higher number of abnormal reflexes (relative risk [RR] = 1.49, 95% CI = 1.12, 1.98), and total DAP levels were also marginally associated with abnormal reflexes (RR = 1.32, 95% CI = 0.99, 1.77), whereas DMP levels were not significantly associated (RR = 1.13, 95% CI = 0.90, 1.41). Detectable MDA levels in maternal prenatal urine were also associated with a significantly higher number of abnormal reflexes (RR = 2.24, 95% CI = 1.55, 3.24). When levels of DAPs, DMPs, and DEPs were categorized into quartiles, some positive associations with number of abnormal reflexes were still detected, but not in a monotonic exposure–response pattern. When the number of abnormal reflexes was dichotomized as ≥ 2 or < 2 and analyses were stratified by infant age, associations with maternal prenatal urinary DAPs, DMPs, and DEPs were stronger for those aged ≥ 2 days, whereas the association with detectable MDA was stronger for those aged 1 day. Statistically significant interactions between maternal prenatal plasma PON1 expression levels and urinary DAP and DMP metabolite levels were detected with risk of ≥ 2 abnormal reflexes as the outcome. Specifically, the RR per-unit increase in prenatal DAPs was 2.38 (95% CI = 1.37, 4.15) for those in the lowest tertile of PON1 expression level versus 0.76 (95% CI = 0.48, 1.20) for those in the highest tertile, and the RR for prenatal DMPs was 1.96 (95% CI = 1.27, 3.03) for those in the lowest tertile of PON1 expression level versus 0.73 (0.56, 0.96) for those in the highest tertile. Associations with prenatal DEPs did not vary significantly by PON1 expression.

In analyses using the Bayley Scales at 12 and 24 months, maternal prenatal urinary DAP and DMP metabolite levels (but not DEP levels) were associated with significantly lower scores on the Mental Development Index at 12 months among blacks and Hispanics (beta per $\log_{10}$-unit increase in DAPs = −3.29, 95% CI = −5.88, −0.70; beta for DMPs = −3.35, 95% CI = −5.64, −1.06), but significantly higher scores among whites (beta for DAPs = 4.77, 95% CI = 0.69, 8.86; beta for DMPs = 4.45, 95% CI = 0.82, 8.08) (Table 2) (Engel et al. 2011). When analyses of the Bayley Mental Development Index at 12 months were stratified by maternal $PON1_{192}$ genotype, interactions were observed among blacks and Hispanics, with significantly lower scores among those carrying the $PON1_{192}$ QR or RR genotype (i.e., heterozygotes and low-activity homozygotes) than QQ homozygotes (e.g., beta per $\log_{10}$-unit increase in DAPs = −4.94, 95% CI = −7.87, −2.07 for $PON1_{192}$ QR/RR carriers; beta = 5.72, 95% CI = −0.48, 11.92 for $PON1_{192}$ QQ carriers). However, no significant interactions were found with the $PON1_{L55M}$ or $PON1_{-108C>T}$ polymorphism, or with PON1 enzymatic activity for any neurodevelopmental outcome assessed. No significant associations were detected between maternal prenatal urinary DAP, DMP, or DEP levels and the Bayley Mental Development Index at 24 months (including after stratification by race/ethnicity or $PON1_{192}$ genotype) or the Bayley Psychomotor Development Index at 12 or 24 months (including after stratification by race/ethnicity). Moreover, DAP, DMP, and DEP levels were not significantly associated with any Wechsler Intelligence Scale measures, including Full-Scale IQ, Perceptual Reasoning, Verbal Comprehension, Processing Speed, and Working Memory (assessed at ages 7–9 years only) at 6, 7–9, or 6–9 years. Only after stratification by $PON1_{192}$ genotype were significant inverse associations detected between maternal prenatal urinary DAP and DMP levels and the Wechsler Perceptual Reasoning Index (e.g., beta per $\log_{10}$-unit increase in DAPs = −0.56, 95% CI = −4.80, 3.68 for $PON1_{192}$ QR/RR carriers; beta = −7.52, 95% CI = −14.53, −0.51 for $PON1_{192}$ QQ carriers). No significant interactions with $PON1_{192}$ genotype were observed for the Wechsler measures of Full-Scale IQ or Verbal Comprehension.

Key strengths and limitations of the Mount Sinai CECS were delineated earlier and apply equally to the analyses of neurological outcomes. Selection bias due to unequal enrollment rates may not have a major influence on associations with long-term childhood neurological outcomes, unless participation varied by strong neurological risk factors, but selection bias due to unequal follow-up is a reasonable concern. For example, 311 (77%) of 404 eligible infants completed the Brazelton Neonatal Behavioral Assessment Scale before hospital discharge, excluding those admitted to the Neonatal Intensive Care Unit (NICU), those delivered and discharged over a weekend, those whose parent refused, those who were not testable, and those for whom study personnel were unavailable; thus, selection bias could have occurred if exclusions were associated with both DAP metabolite levels and neonatal behavioral outcomes. Multiple comparisons potentially leading to chance findings are a particular concern in these analyses, given the large number of outcomes and subgroups examined, along with the apparent lack of *a priori* hypotheses regarding why some but not other neurological outcomes might be associated with OP metabolites, or why associations might be observed in some but not other subgroups by age and race/ethnicity. Consequently, although the associations of maternal prenatal urinary levels of DAP and DEP metabolites and detectable MDA with abnormal neonatal reflexes were noteworthy, the absence of a monotonic exposure–response pattern, along with the absence of association with motor performance, autonomic stability, and other neurological outcomes, detracts from the coherence of these findings. Likewise, the persuasiveness of the inverse associations of maternal prenatal urinary levels of DAP and DMP metabolites with mental development at 12 months in blacks and Hispanics, especially in $PON1_{192}$ QR/RR carriers, is undermined by the positive associations in whites, the absence of any association at 24 months, and the lack of any interaction with other *PON1* genotypes or PON1 activity levels. The stronger inverse association of prenatal DAP and DMP levels with perceptual reasoning in 6- to 9-year-olds in $PON1_{192}$ QQ carriers also runs counter to expectation. Consequently, the few observed significant associations among a large number of statistically null associations, without a discernable pattern, cannot reliably be interpreted as causal, and require confirmation in independent studies.

Center for the Health Assessment of Mothers and Children of Salinas

The basic methods of the CHAMACOS birth cohort study were described earlier; follow-up for neurodevelopmental outcomes continued through age 7 years (Table 1) (Bouchard et al. 2011, Eskenazi et al. 2010, Eskenazi et al. 2007, Marks et al. 2010, Quiros-Alcala et al. 2011, Young et al. 2005). Geometric mean urinary DAP metabolite levels measured in children increased with age: 45.5 nmol/L (95% CI = 39.6, 52.3) at 6 months, 59.5 = nmol/L (51.7, 68.5) at 12 months, 70.9 nmol/L (61.4, 81.9) at 24 months, 77.5 nmol/L (65.4, 91.9) at 3.5 years, and 92.6 nmol/L (78.6, 109.0) at 5 years (Eskenazi et al. 2007, Marks et al. 2010). Neurodevelopmental outcomes were measured using the Brazelton Neonatal Behavioral Assessment Scale administered by 62 days (2 months); the Bayley Scales of Infant Development, 2nd Edition, administered at 6, 12, and 24 months; an autonomic nervous system reactivity protocol that measured heart rate, respiratory sinus arrhythmia, and pre-ejection period following social, physical, and emotional challenges (and cognitive challenges for older children) administered at 6 months and 1, 3.5, and 5 years; the mother-completed Child Behavior Checklist for ages 1.5–5 years administered at 2, 3.5, and 5 years; the NEPSY® visual attention subtest, 2nd Edition, administered at 3.5 years; the Conners' Kiddie Continuous Performance Test, which assesses reaction time, accuracy, and impulse control for ADHD screening using an interactive computer program; the Hillside Behavior Rating Scale, which assesses motor activity and distractibility for ADHD screening, administered at 5 years; and the Wechsler Intelligence Scale for Children, 4th Edition, administered at 7 years.

Among infants assessed at or before age 2 months, no significant association was observed between maternal prenatal average urinary levels of DAPs, DMPs, or DEPs and the Brazelton habituation, orientation, motor performance, range of state, or regulation of state cluster, either overall or among neonates assessed at age ≤ 3 days or > 3 days (Table 2) (Young et al. 2005). Maternal prenatal urinary DEP metabolite levels, but not DAPs or DMPs, were significantly associated with a higher score on the autonomic stability cluster, which includes tremors, startles, and skin color, at age ≤ 3 days (beta per $\log_{10}$-unit increase = 0.31, 95% CI = 0.01, 0.61), but not at age > 3 days (beta = − 0.16, 95% CI = − 0.47, 0.14) or overall. By contrast, maternal prenatal urinary DAP, DMP, and DEP metabolite levels were all significantly associated with a higher number of abnormal reflexes, especially at age > 3 days (beta for DAPs = 0.53, 95% CI = 0.23, 0.82; beta for DMPs = 0.41, 95% CI = 0.12, 0.69; beta for DEPs = 0.37, 95% CI = 0.09, 0.64), but not at age ≤ 3 days (beta for DAPs = − 0.01, 95% CI = − 0.24, 0.22; beta for DMPs = − 0.00, 95% CI = − 0.21, 0.20; beta for DEPs = 0.08, 95% CI = − 0.16, 0.32). When the number of abnormal reflexes was dichotomized at > 3 versus ≤ 3, maternal prenatal urinary DAP, DMP, and DEP levels were categorized into quintiles, and the analysis was restricted to neonates aged > 3 days at assessment, statistically significant positive exposure–response trends were observed for each metabolite type. The OR for > 3 abnormal reflexes per $\log_{10}$-unit increase in metabolite concentration was 4.9 (95% CI = 1.5, 16.1) for DAPs, 3.2 (95% CI = 1.1, 9.8) for DMPs, and 3.4 (95% CI = 1.2, 9.9) for DEPs. No associations with any neonatal neurodevelopmental outcome were detected with maternal post-delivery urinary metabolite levels.

A significant inverse association was detected between maternal prenatal urinary DAP and DMP levels and the Bayley Mental Development Index at 24 months (beta per $\log_{10}$-unit increase in DAPs = − 3.54, 95% CI = − 6.59, − 0.49; beta for DMPs = − 3.64, 95% CI = − 6.36, − 0.91) (Table 2) (Eskenazi et al. 2007). By contrast, child urinary DAP and DMP levels at 24 months were positively associated with the Mental Development Index (beta for DAPs = 2.37, 95% CI = 0.50, 4.24; beta for DMPs = 2.01, 95% CI = 0.24, 3.78). Child urinary DEP levels at 12 months were also positively associated with the Mental Development Index at that age (beta = 1.89, 95% CI = 0.21, 3.58). Otherwise, associations of maternal prenatal and child urinary DAP, DMP, and DEP metabolites, as well as maternal prenatal MDA and TCPy levels, with the Mental Development Index at 6, 12, and 24 months were statistically non-significant, and all associations with the Psychomotor Development Index at those ages were non-significant. Maternal prenatal and 24-month child urinary levels of DAPs, DMPs, DEPs, MDA, and TCPy were not significantly associated with a clinically borderline score (> 93rd percentile) for attention problems or ADHD as assessed by the Child Behavior Checklist at 24 months. However, maternal prenatal urinary levels of DAPs and DMPs were at least marginally significantly associated with a higher odds of clinical pervasive developmental disorder (> 97th percentile) as assessed by the Child Behavior Checklist at 24 months (OR for DAPs = 2.25, 95% CI = 0.99, 5.16; OR for DMPs = 2.19, 95% CI = 1.05, 4.58; OR for DEPs = 0.88, 0.37, 2.07), as were all three types of metabolites in children (OR for DAPs = 1.71, 95% CI = 1.02, 2.87; OR for DMPs = 1.52, 95% CI = 0.94, 2.45; OR for DEPs = 1.72, 1.12, 2.64). Maternal prenatal urinary MDA and TCPy levels were not significantly associated with pervasive developmental disorder at 24 months.

When associations between maternal prenatal urinary DAP, DMP, and DEP levels and the Bayley Mental and Psychomotor Development Indices and Child Behavior Checklist pervasive developmental disorder score were stratified by child or maternal *PON1* genotype, or by umbilical cord or maternal blood PON1 activity or quantity, no statistically significant interactions were detected, and stronger associations were not consistently detected among those with lower-activity genotypes (i.e., $PON1_{192}$ RR and $PON1_{-108}$ TT) or lower enzyme levels (Table 2) (Eskenazi et al. 2010).

At age 3.5 years, maternal prenatal urinary DAP, DMP, and DEP levels were not significantly associated with attention problems or ADHD as assessed by the Child Behavior Checklist, whether the outcomes were analyzed as continuous or categorical variables dichotomized at clinically borderline scores (> 93rd percentile) (Table 2) (Marks et al. 2010). However, several OR point estimates were around 3.0, with wide 95% CIs due to the small number of borderline scores at that age. Prenatal urinary DAP, DMP, and DEP metabolite levels were also unassociated with the NEPSY-II visual attention score at age 3.5 years. Associations with Child Behavior Checklist measures of attention problems and ADHD at age 5 years were attenuated and statistically non-significant in analyses

of dichotomized scores at age 5 years, but significant positive associations were detected with scores analyzed as continuous outcomes (beta per 10-fold increase in DAPs = 0.7, 95% CI = 0.2, 1.2 for attention problems; beta = 1.3, 95% CI = 0.4, 2.1 for ADHD). Whereas no significant associations were detected between maternal prenatal urinary DAP, DMP, or DEP metabolite levels and markedly atypical scores for omissions, commissions, or hit reaction time on the Conners' Kiddie Continuous Performance Test or the ADHD Confidence Index analyzed as the continuous variable at 5 years, the odds of having an ADHD Confidence Index above the 70th percentile (OR per 10-fold increase in DAPs = 5.1, 95% CI = 1.7, 15.7), a Hillside Behavior Rating Scale attention problems score ≥7 out of 12 (OR for DAPs = 3.0, 95% CI = 0.9, 9.8), or a positive composite ADHD indicator (OR for DAPs = 3.5, 95% CI = 1.1, 10.7) were all at least marginally significantly increased in association with higher prenatal metabolite concentrations. Some heterogeneity was detected by sex, with boys generally showing stronger associations than girls. Associations with child urinary OP metabolite levels were weaker and not statistically significant.

At age 7 years, significant inverse associations were found between maternal prenatal urinary levels of DAPs, DMPs, and DEPs and Wechsler Intelligence Scale measures of Working Memory (e.g., beta per $\log_{10}$-unit increase in DAPs averaged from the first and second halves of pregnancy = −4.3, 95% CI = −7.7, −0.9), Processing Speed (beta for averaged DAPs = −3.4, 95% CI = −6.8, −0.1), Verbal Comprehension (beta for averaged DAPs = −5.3, 95% CI = −8.6, −2.0), Perceptual Reasoning (beta for averaged DAPs = −4.0, 95% CI = −7.9, −0.1), and Full-Scale IQ (beta for averaged DAPs = −5.6, 95% CI −9.0, −2.2) (Table 2) (Bouchard et al. 2011). When maternal averaged prenatal urinary DAP levels were categorized by quintile, inverse exposure–response trends were observed for all five outcomes, with an average difference of 7.0 Full-Scale IQ points between the highest and lowest quintiles of prenatal DAPs. Estimates of association did not differ substantially among maternal early prenatal, late prenatal, and postnatal urinary DAP concentrations, nor were marked changes observed after additional adjustment for other environmental contaminants, standardization by creatinine, stratification by sex, or restriction to Spanish-speaking children. However, child urinary DAP levels at 6, 12, 24, 42, or 60 months, or at all ages taking the area under the concentration–time curve, were not significantly associated with any of the Wechsler Intelligence Scale measures at age 7 years.

Associations of maternal prenatal and child urinary DAPs, DMPs, and DEPs with both resting and reactivity measures of respiratory sinus arrhythmia, heart rate, and pre-ejection period were tested at ages 6 months, 1 year, 3.5 years, and 5 years (Table 2) (Quiros-Alcala et al. 2011). In addition, cumulative measures of prenatal (14-week and 26-week) and childhood (6 months to 5 years, based on area under the concentration–time curve calculations) urinary DAP, DMP, and DEP metabolite levels were analyzed with respect to resting and reactivity measures at age 5 years. Among the numerous associations tested, significant associations were found only between child DAPs and DMPs and resting respiratory sinus arrhythmia score at 6 months (beta per $\log_{10}$-unit increase in DAPs = −0.27, 95% CI = −0.48, −0.06; beta for DMPs = −0.24, 95% CI = −0.42, −0.05; beta for); between maternal prenatal DMPs and child DEPs and resting pre-ejection period at 1 year (beta for prenatal DMPs = 3.77 milliseconds, 95% CI = 0.21, 7.33; beta for child DEPs = 4.33 milliseconds, 95% CI = 1.24, 7.42); between maternal prenatal DMPs and reactive pre-ejection period at 6 months (beta = 1.2 milliseconds, 95% CI = 0.03, 2.40); between maternal prenatal DAPs and DMPs and reactive respiratory sinus arrhythmia at 1 year (beta for DAPs = 0.24, 95% CI = 0.03, 0.46; beta for DMPs = 0.25, 95% CI = 0.05, 0.45); and between cumulative maternal prenatal DEPs and resting heart rate (beta = −3.19 beats per minute, 95% CI = −6.29, −0.09). Otherwise, all tested associations were statistically non-significant, and estimated coefficients showed no consistent direction of association. When basic measures of autonomic nervous system function at 6 months, 1 year, 3.5 years, and 5 years were combined into four profiles (coactivation of both sympathetic and parasympathetic nervous systems; coinhibition of both nervous systems; reciprocal activation of parasympathetic and inhibition of sympathetic nervous systems; or reciprocal activation of sympathetic and inhibition of parasympathetic nervous systems), no significant differences in geometric mean urinary DAP concentrations were found based on maternal prenatal or child specimens, nor were these profiles associated with consistently high versus low urinary DAP metabolite levels in gestation, childhood, or both.

As with other prospective birth cohort studies described earlier in the section on birth outcomes, the main strengths and limitations of the CHAMACOS study were discussed above, with perhaps a lower probability of selection bias with respect to enrollment rates but a higher probability in terms of follow-up rates, and greater concerns about multiple comparisons due to the larger number of neurodevelopmental risk factors assessed. Overall, the neurodevelopmental results from CHAMACOS varied by outcome, metabolite type, age group, and timing of exposure assessment. The associations of maternal prenatal urinary DAPs, DMPs, and DEPs with a higher number of abnormal reflexes at age >3 days were fairly consistent across metabolites and outcome classifications, but they should be balanced against the null associations with abnormal reflexes at ≤3 days and with other neonatal behavioral outcomes (except autonomic stability at <3 days, which was positively associated with maternal prenatal urinary DEP levels). Another noteworthy finding is the positive association of maternal prenatal DAP and DMP levels and child DAP, DMP, and DEP levels (but not MDA or TCPy levels) with the risk of pervasive developmental disorder score above the clinical cutoff at 24 months. These results are notable and warrant further evaluation, although the reliance on mother-reported symptoms to classify this outcome leaves open the possibility of misclassification, whether non-differential or differential by exposure status. The inverse associations of prenatal DAP and DMP levels with the Bayley Mental Development Index at 24 months are less compelling, given the opposite associations with child metabolite levels and the statistically null associations with prenatal DEP, MDA, and TCPy levels as exposures, and with the Mental Development Index at 6 and 12 months as outcomes. The results for autonomic nervous system function were consistently null.

Two sets of striking associations with adverse neurodevelopmental outcomes were reported in CHAMACOS. The first were the positive associations between prenatal DAP and DMP levels and continuous Child Behavior Checklist scores for attention problems and ADHD and dichotomized indicators for ADHD Confidence Index, Hillside Behavioral Rating Scale attention problems, and a composite ADHD indicator at 5 years. Significant associations were not detected with dichotomized Child Behavior Checklist scores for attention problems and ADHD, dichotomized atypical scores on the Conners' Kiddie Continuous Performance Test, and continuous ADHD Confidence Index, although point estimates were generally in the positive direction. However, associations were attenuated and mostly non-significant for child urinary DAP, DMP, and DEP levels. The somewhat inconsistent findings raise the question of whether some measures are more valid than others for capturing ADHD risk, and the heterogeneity of associations between boys and girls—with some inverse and mostly nearly null point estimates among girls—is not readily explained by known biological mechanisms. The other salient results in CHAMACOS were the inverse associations of maternal prenatal (but not child) urinary DAP, DMP, and DEP levels with all five Wechsler Intelligence Scales at age 7 years. The consistency of these findings is unlikely to be due to chance. The methodological limitations of this study, especially with regard to OP insecticide exposure assessment, prevent a causal interpretation of these findings, but the robust associations with impaired behavioral and cognitive development in school-aged children in CHAMACOS warrant attention and replication in independent studies.

Health Outcomes and Measures of the Environment Study

The HOME Study, described earlier, included 350 mothers who provided urine specimens at 16 ± 4 and 26 ± 4 weeks of gestation, and whose infants completed the NICU Network Neurobehavioral Scale at home at approximately 5 weeks of age (Table 1) (Yolton et al. 2013). The scale covers 13 dimensions: habituation (excluded from analysis because it was omitted for sleeping infants), attention, arousal, self-regulation, need for special handling from the examiner, quality of movement, excitability, lethargy, non-optimal reflexes, asymmetrical reflexes, hypertonicity, hypotonicity, and stress/abstinence. In multivariate regression models, significant associations were detected between creatinine-standardized maternal prenatal urinary DEP levels averaged over 16 and 26 weeks and increased attention (beta per $\log_2$-unit increase = 0.066, SE = 0.033); between DEP levels at 16 weeks and decreased lethargy (beta = −0.069, SE = 0.034) and decreased hypotonia (beta = −0.101, SE = 0.045, with hypotonia dichotomized as none vs. any); and between DAP levels at 16 weeks and decreased autonomic stress (beta = −0.010, SE = 0.004) (Table 2). No other significant associations were detected between maternal prenatal urinary DAPs, DMPs, or DEPs at 16 weeks, 26 weeks, or the average of the two, and any of the other eight dimensions assessed. In secondary analyses, latent profile analysis was used to group infants together based on NICU Network Neurobehavioral Scale scores into one of three patterns: social/easy-going ($n = 157$), hypotonic ($n = 110$), or high-arousal/difficult ($n = 83$). A significantly decreased odds of being classified as hypotonic, compared with social/easy-going, was detected among infants whose mothers had higher creatinine-standardized urinary DEP levels at 16 weeks (OR per $\log_2$-unit increase = 0.89, 95% CI = 0.81, 0.99). Otherwise, no significant associations were observed between maternal prenatal urinary DAPs, DMPs, or DEPs at any time point and the odds of being classified as hypotonic or high-arousal/difficult, although several borderline significant associations were found in both directions, with no apparent consistency by exposure or outcome.

The main strengths and limitations of the HOME Study were discussed earlier and previous comments also apply to this analysis. The use of only three profiles to classify neurobehavior in secondary analyses may be an oversimplification of a complex neurobehavioral scale (Lester et al. 2004). Although several results were generally in the same direction, with higher maternal prenatal urinary DAP or DEP levels being associated with better neurobehavioral outcomes (i.e., increased attention, decreased lethargy, decreased hypotonia, and decreased autonomic stress), these significant associations were selected among many others that were tested and found to be null. Thus, these findings cannot reliably be interpreted as demonstrating a beneficial causal effect of prenatal OP insecticide exposure on behavioral neurodevelopment in infants.

Children Pesticide Survey

From the Children Pesticide Survey, a cross-sectional study of children living in an agricultural community in southern Arizona in 1998–2000, a subgroup of 25 school-aged children was selected for analysis based on detectable DAP levels (≥ 25 µg/mL; metabolite not specified) in an initial urine sample, and 23 other children were selected who had undetectable levels (Table 1) (Lizardi et al. 2008). Subsequently, urinary DAPs were re-measured in a first-void urine sample, and a cognitive assessment was conducted on the same day using a short form of the Wechsler Intelligence Scale for Children Third Edition, the Children's Memory Scale, the Wisconsin Card Sorting Test, and the Trail Making Test A and B. In addition, the Child Behavior Checklist 4–18 and the Teacher Report Form were used to assess behavioral outcomes. Based on the urine samples collected on the day of the cognitive assessment, all 48 children had detectable levels of DMP, although average levels remained significantly higher in the originally designated "exposed" group (mean = 110 µg/L, 95% CI = 83, 139) than in the originally designated "unexposed" group (mean = 49 µg/L, 95% CI = 36, 63) after excluding one outlier from each group (519 µg/L in the "exposed" group and 850 µg/L in the "unexposed" group).

Although children in the "exposed" group took significantly longer time (mean = 283 seconds, 95% CI = 224, 341) to complete the Trail Making Test B than children in the "unexposed" group (mean = 204 seconds, 95% CI = 172, 236), none of the other cognitive or behavioral measures differed significantly between the groups in unadjusted analyses, excluding the two outliers (Table 2). Concurrent urinary DAP levels (analyzed as the sum of all six metabolites) were modestly and statistically significantly correlated with some measures of the Wisconsin Card Sorting Test ($\rho = 0.31$–0.38, $P \leq 0.03$), but not after exclusion of the two outliers. Moreover, no significant correla-

tions were detected with the other cognitive measures, including the Wechsler Intelligence Scale, the Children's Memory Scale, and both Trail Making Tests. Correlations between concurrent urinary DAP levels and behavioral measures were not estimated.

A key limitation of this study is its cross-sectional design: because exposures and outcomes were measured on the same day, a cause-and-effect relationship cannot be established. Reverse causality due to an influence of childhood behavior on diet, as the major source of OP exposure, is plausible. Even without such an effect, it seems unlikely that DAP metabolites are etiologically relevant to cognitive performance measured on the same day. Other limitations include the lack of adjustment for any confounders, the small study size (resulting in unstable estimates and, possibly, insufficient statistical power to detect any associations), the large number of outcomes tested (resulting in the expectation of several chance findings), and the use of a single sample of urinary DAP levels. In particular, the fact that the originally designated "unexposed" group had detectable urinary DAP levels at the second assessment underscores the intra-individual variability of these metabolites. Participation rates were not reported, precluding an assessment of potential selection bias. These predominantly null results add little insight into possible adverse neurodevelopmental effects of exposure to OP insecticides.

National Health and Nutrition Examination Survey

The NHANES is a continuous series of population-based health surveys designed to assess the health and nutritional status of approximately 5000 representative, non-institutionalized adults and children in the United States each year (Table 1) (Bouchard et al. 2010). In 2000–2004, NHANES data on six urinary DAP metabolites and ADHD were available for 1,139 children (119 with ADHD) aged 8–15 years, where ADHD diagnostic status during the previous year was assessed based on symptoms reported by the mother or another caretaker in a telephone interview using the Diagnostic Interview Schedule for Children IV (using slightly modified criteria from the DSM-IV), or based on reported use of ADHD medication. Geometric mean urinary levels, which were measured 2–3 weeks before the interview, were 68.3 nmol/L (IQR = 24.4–186.0) for DAPs, 41.3 nmol/L (IQR = 10.1–130.7) for DMPs, and 11.0 nmol/L (IQR = 2.1–35.0) for DEPs.

A 10-fold increase in urinary DAP or DMP metabolite levels was associated with a significantly increased prevalence of ADHD as defined based on diagnostic interview criteria or ADHD medication use (adjusted OR = 1.35, 95% CI = 1.10, 1.67 for DAPs; OR = 1.72, 95% CI = 1.31, 2.28 for DMPs), or based on diagnostic interview criteria alone (Table 2). A positive exposure–response gradient was observed across undetectable, below-median, and above-median urinary levels of DMPs. Urinary DEP metabolite levels were not significantly associated with prevalent ADHD (OR = 0.80, 95% CI = 0.60, 1.05). However, urinary DAPs, DMPs, and DEPs were all significantly associated with a higher prevalence of the hyperactive/impulsive subtype of ADHD ($n = 21$ children; OR per 10-fold increase = 1.85, 95% CI = 1.04, 3.27 for DAPs; OR = 2.13, 95% CI = 1.08, 4.20 for DMPs; OR = 2.15, 95% CI = 1.06, 4.40 for DEPs), whereas only DMPs were marginally significantly associated with the inattentive subtype of ADHD ($n = 69$ children; OR = 1.47, 95% CI = 0.99, 2.19) and no metabolites were significantly associated with the combined hyperactive/impulsive and inattentive subtype of ADHD ($n = 29$ children; OR = 1.30, 95% CI = 0.48, 3.48 for DMPs).

This study is strengthened by its population-based sample selection and by the availability of detailed interview and physical examination data to adjust for potential confounders (albeit not diet [Millichap and Yee 2012]).

Major methodological limitations are the cross-sectional design and the measurement of urinary DAPs at a single point in time. Outcomes were classified based on parent- or caretaker-reported symptoms, which could have been differentially misclassified if, for example, accuracy of reporting varied by dietary patterns or other lifestyle characteristics related to OP insecticide exposure. As mentioned by the authors, the observed associations might be due to reverse causality—i.e., ADHD-related behaviors, such as dietary changes (Millichap and Yee 2012)—that could result in higher exposure to OP insecticides and their metabolites. Participation rates among ADHD and non-ADHD children were unknown, given that ADHD diagnosis was predicated on participation; therefore, the potential for selection bias could not be assessed. The authors did not suggest a mechanism to explain the stronger associations of DAP metabolites with the hyperactive/impulsive subtype of ADHD than others. Overall, the results of this study indicate a positive association between DAP metabolite levels and the prevalence of ADHD, but causal inference about the effects of OP insecticide exposure is limited by the cross-sectional study design.

Shanghai cross-sectional study

In a cross-sectional study of 301 healthy 2-year-olds recruited in 2008 from two community hospitals in Shanghai, urinary levels of five DAP metabolites were measured in spot urine samples on the same day on which a neurological assessment of motor behavior, adaptive behavior, language behavior, and personal and social behavior was completed using the Gesell Developmental Schedules for 0- to 3-year-old children (Table 1) (Guodong et al. 2012). Geometric mean urinary levels were 2.52 µg/L (IQR = < 2.0 [LOD]–3.41) for DMP, 1.56 µg/L (IQR = < 1.0–1.63) for DMTP, 1.78 µg/L (IQR = < 1.0–2.89) for DEP, and 3.18 µg/L (IQR = < 1.0–7.26) for DETP; DEDTP was detected in only 2.7% of subjects. No significant associations were observed between a $\log_{10}$-unit increase in creatinine-adjusted urinary DAPs, DMPs, or DEPs and any of the four Gesell Developmental Schedule scores, and estimated coefficients were not consistently above or below zero across metabolites or outcome measures (Table 2).

The high participation rate in this study (97%) minimizes selection bias, but the cross-sectional design and reliance on a single biospecimen remain the major limitations. Information on confounders was somewhat limited, although several covariates were included in multivariate models, and the direction and magnitude of any uncontrolled confounding are unpredictable. Again, the lack of

prospectively collected serial OP metabolites prevents this study from fully assessing the associations between exposure to OP insecticides and neurodevelopmental behavioral outcomes in young children.

Canadian health measures survey

In the first cycle (2007–2009) of the cross-sectional Canadian Health Measures Survey, the Canadian counterpart to NHANES, 1081 children aged 6–11 years were enrolled, including 1030 (95%) with spot urine measurements of six DAP metabolites measured within two weeks of parental completion of the Strengths and Difficulties Questionnaire to assess mental and behavioral outcomes (Table 1) (Oulhote and Bouchard 2013). The five-dimension scales of this questionnaire evaluate emotional symptoms, conduct problems, hyperactivity/inattention, peer problems, and prosocial behavior (not analyzed due to insufficient variability), each scored on a 10-point scale; a global total difficulties scale is computed based on the sum of all scales except prosocial behavior. Scores were dichotomized between high and low/borderline using cutoffs recommended by the author of the instrument. Of the eligible children, 779 (72%) had complete covariate data and were included in the analysis. The median urinary level of DAPs was 99.2 nmol/L (IQR = 34.3–273.3), that of DMPs was 62.0 (IQR = 18.7–192.8), and that of DEPs was 25.0 (IQR = 10.5–51.3).

When analyzed on the $\log_{10}$ scale and adjusted for multiple covariates, with or without creatinine standardization, urinary DAPs, DMPs, and DEPs were all statistically unassociated with elevated scores for total difficulties (OR for DAPs = 0.6, 95% CI = 0.3, 1.3; OR for DMPs = 0.8, 95% CI = 0.4, 1.6; OR for DEPs = 0.3, 95% CI = 0.1, 1.8), conduct problems, emotional symptoms, hyperactivity/inattention, and peer problems (Table 2). No significant heterogeneity was observed by child sex.

The methodological strengths and limitations of this study are essentially the same as those of the NHANES study described above (Bouchard et al. 2010). Advantages include the population-based setting and extensive information on potential confounders (but not diet), whereas major drawbacks include the cross-sectional design and the one-time spot urine measurement of DAP metabolites. Parent-reported outcome measures were subject to misclassification that might have been differential. Selection bias could have influenced the results in unpredictable ways if participation in the Canadian Health Measures Survey or provision of complete covariate data were related to both the exposure and the outcome. In light of these limitations and the statistically null findings, this study offers no evidence to support a causal effect of OP insecticide exposure on behavioral problems in children.

Early life Exposed in Mexico to Environmental Toxicants Study

The Early Life Exposed in Mexico to Environmental Toxicants (ELEMENT) study sequentially enrolled 827 healthy pregnant women from a general hospital and affiliated clinics in Mexico City (Table 1) (Fortenberry et al. 2014). Of the original cohort participants, 187 (23%) mother–child pairs had third-trimester urine specimens and completed child psychometric assessments to screen for ADHD-related symptoms at ages 6–11 years in 2007–2011; these assessments included the Conners' Parental Rating Scales-Revised, the Behavior Assessment System for Children–Parental Rating Scales, and Conners' Continuous Performance Test. The Conners' Parental Rating Scales-Revised, a parent-completed assessment tool for children and adolescents aged 3–17 years, included scales for an ADHD index, global restlessness/impulsivity, hyperactivity/impulsivity ADHD, inattention ADHD, and combined-type ADHD, mostly based on guidelines from the DSM-IV. The Behavioral Assessment System for Children–Parental Rating Scales, a parent-completed assessment tool for children aged 6–11 years, were used to assess attention problems and hyperactivity. The geometric mean concentration of TCPy in maternal prenatal urine was 1.76 ng/mL (IQR = 0.91–3.57). In a subset of 21 subjects who provided prenatal urine specimens in all three trimesters, the geometric mean concentration did not vary significantly across trimesters, but significant within-subject variability was detected (intraclass correlation = 0.41 without correction for specific gravity and 0.29 with correction).

No significant association, including after stratification by sex, was found between maternal prenatal urinary TCPy level and any of the outcome measures studied, including all ADHD and restlessness/impulsivity scales based on the Conners' Parental Rating Scales-Revised; the two scales for attention problems and hyperactivity based on the Behavior Assessment System for Children–Parental Rating Scales; and the clinical index for ADHD and the hit reaction time block change measure (used to assess vigilance or sustained attention) based on the Conners' Continuous Performance Test (Table 2). The authors highlighted "suggestive trends" (with P values >0.05 but <0.10) between maternal prenatal urinary TCPy and increasing hit reaction time block change and the Conners' Parental ADHD index among boys, but no apparent trends were detected ($P = 0.18$–0.99) for any other ADHD screening measures.

The ELEMENT study benefits from prospective collection of prenatal urine specimens, its measurement of a specific metabolite of chlorpyrifos, and its adjustment for numerous potential confounders (excluding diet). The scope of the study is confined by the measurement of only one OP insecticide metabolite. Other limitations, which are shared by other studies discussed in this review, include the lack of serial biomonitoring, potential selection bias, possible outcome misclassification due to the use of parent-reported data, a modest number of subjects, and multiple comparisons, with no *a priori* hypothesis regarding why TCPy should be associated with some measures of ADHD-related symptoms but not others. Thus, chance must be considered as a reasonable explanation for the two marginally significant trends detected among at least 27 tested. In general, the results of this study suggest no consistent or convincing associations between prenatal TCPy levels and ADHD-related symptoms.

Shenyang birth cohort

In another prospective birth cohort study, 249 healthy pregnant women were enrolled from a hospital in Shenyang, China, between 2011 and 2012 and followed through delivery of a healthy neonate (Table 1) (Zhang et al. 2014). The Neonatal

Behavioral Neurological Assessment, developed for Chinese newborns, was administered at 3 days of age to measure functional abilities, reflexes and responses, and behavioral status based on five scales: behavior, passive tone, active tone, primary reflexes, and general assessment. Concentrations of five DAP metabolites were measured in prenatal maternal urine (timing of collection not specified), with the following geometric means: 18.03 μg/L (IQR = 7.83–39.43) for DMP, 8.53 μg/L (IQR = 3.4–15.67) for DMTP, 7.14 μg/L (IQR = 3.54–17.17) for DEP, 5.64 μg/L (IQR = 2.34–13.55) for DETP, and <1 μg/L (LOD; IQR = LOD–LOD) for DEDTP.

In adjusted linear regression models, $\log_{10}$-unit increases in maternal prenatal urinary levels of DAPs, DMPs, and DEPs were all significantly associated with lower summary scores on the Neonatal Behavioral Neurological Assessment (beta for DAPs = −1.78, 95% CI = −2.12, −1.45; beta for DMPs = −0.96, 95% CI = −1.35, −0.57; beta for DEPs = −0.88, 95% CI = −1.30, −0.47) (Table 2). Significant inverse associations were also observed between maternal prenatal DAPs and DEPs and behavior, between DAPs and DMPs and passive tone, between DAPs and DMPs and active tone, and between DAPs and DMPs and primary reflexes. These associations were detected in both boys and girls, and with and without creatinine standardization. Analyses with maternal prenatal urinary DAP concentrations categorized into quintiles were consistent with linear inverse exposure–response associations with all five outcomes examined. When regression coefficients were standardized, the associations between maternal prenatal urinary DAP levels and all outcomes were stronger than those with gestational age, cord blood lead levels, and maternal prenatal BMI.

Like other birth cohort studies, the Shenyang study is strengthened by the measurement of urinary DAP metabolite levels prior to the measurement of neurological outcomes, which rules out reverse causality. However, urine specimens appear to have been collected from various subjects at different times throughout pregnancy, and it may or may not be plausible that exposures at different stages of neurodevelopment would have the same effect on behavioral outcomes. The participation rate (81%) among eligible women was relatively high, thereby reducing concerns about selection bias, and information was collected on numerous potential confounders, thereby lessening the probability of strong confounding. However, due to the reliance on a single biospecimen and the measurement of non-specific DAP metabolites, the observed inverse associations between prenatal DAP metabolite levels and neonatal behavioral outcomes cannot reliably be interpreted as causal. In addition, the applicability of the outcome assessment instrument outside of China, where it was developed and tested, is unknown.

Bradford Hill evaluation of weight of evidence

Some measures of neurodevelopmental outcomes, including the Brazelton Neonatal Behavioral Assessment Scale, the Bayley Scales of Infant Development, the Wechsler Intelligence Scales, the Conners' Parent Rating Scales, and the Child Behavior Checklist, were used in more than one study, but several were not. Although all relevant studies of OP metabolites and neurodevelopmental outcomes were described in the preceding section, outcomes that were uniquely evaluated in only one study (e.g., brain morphology (Rauh et al. 2012) and specific autonomic nervous system functions (Quiros-Alcala et al. 2011)) are not included in the weight-of-evidence evaluation because of the absence of independent results for comparison.

Strength. As in the case of associations with birth outcomes, the strength of observed associations between OP metabolites and neurodevelopmental outcomes cannot be compared readily across studies, due to variations in the unit of exposure measurement, the inconsistent use of logarithmic transformation or creatinine standardization of metabolite levels, outcome measurement and classification methods, and the format of reported results. "Strong" versus "weak" associations also are not objectively defined, especially for continuous exposures and outcomes. Even so, most observed associations entail relatively modest changes in outcomes—for example, ORs and RRs between 0.5 and 2.0, and increases or decrements of a few points on a scale standardized to a mean of 100 and SD of 15. Confounding and bias cannot confidently be ruled out as explanations for associations of such a magnitude. Several ORs around or above 5.0 were reported in the CHAMACOS study (Eskenazi et al. 2010, Marks et al. 2010, Rauh et al. 2006, Young et al. 2005), but most of these were statistically unstable, with lower 95% confidence limits near or below 1.0. Although these associations with large ORs merit a closer look, most of these and other reported associations are statistically non-significant, making them consistent with no association between OP metabolites and neurodevelopmental outcomes.

Consistency. To evaluate the consistency of findings across study settings, we assume that neurodevelopmental outcomes evaluated using different assessment tools are reasonably comparable. In four studies conducted in four different settings, neonatal behavior was evaluated using the Brazelton Neonatal Behavioral Assessment Scale, the NICU Network Neurobehavioral Scale, and the Neonatal Behavioral Neurological Assessment (Engel et al. 2007, Yolton et al. 2013, Young et al. 2005, Zhang et al. 2014). Three of these four studies found an association between prenatal OP metabolite levels and poorer reflexes at or shortly after birth (Engel et al. 2007, Young et al. 2005, Zhang et al. 2014). Three studies also showed no association with any other adverse neonatal behavioral outcomes (Engel et al. 2007, Yolton et al. 2013, Young et al. 2005). The statistically null results for poorer reflexes in the HOME Study (Yolton et al. 2013), in which newborns were older at the time of assessment than those in the other three studies, may suggest that the association is no longer detectable by age 5 weeks. Alternatively, the heterogeneity might be due to chance, confounding or bias, or true differences in study populations or assessment tools.

Among infants and toddlers evaluated in four different study settings, behavioral outcomes were measured using the Bayley Scales of Infant Development and the Gesell Developmental Schedules (Engel et al. 2011, Eskenazi et al. 2010, Eskenazi et al. 2007, Guodong et al. 2012, Lovasi et al. 2011, Rauh et al. 2006). Although all three studies that measured pre- or perinatal OP metabolites and used the Bayley Scales of Infant Development found a significant inverse association between

OP metabolite levels and scores on the Mental Development Index (Engel et al. 2011, Eskenazi et al. 2007, Rauh et al. 2006), this apparent consistency is no longer evident after a closer examination of results. Specifically, the CCCEH study detected an association at 36 months among African American children but not at 12 or 24 months or in Dominican children (Rauh et al. 2006); the Mount Sinai CECS detected an association at 12 months but not at 24 months among black and Hispanic children, and an association in the opposite direction at 12 months among white children (Engel et al. 2011); and the CHAMACOS cohort found an association at 24 months but not at 6 or 12 months (Eskenazi et al. 2007). Thus, none of these studies detected persistent decrements in the Mental Development Index related to OP insecticide exposure across infancy and early childhood age groups. No adverse cross-sectional associations between child urinary OP metabolite levels and mental development at 24 months were reported in CHAMACOS (Eskenazi et al. 2007) and the Shanghai study (Guodong et al. 2012), and most (three out of four) studies did not detect any significant associations with infant psychomotor development (Engel et al. 2011, Eskenazi et al. 2007, Guodong et al. 2012).

Four studies in four separate settings assessed cognitive outcomes in preschool- and school-aged children using the Wechsler Intelligence Scales, the Children's Memory Scale, the Wisconsin Card Sorting Test, and the Trail Making Tests, although only the Wechsler Scales were used in more than one study (Bouchard et al. 2011, Engel et al. 2011, Lizardi et al. 2008, Rauh et al. 2011). Two studies found an inverse association between prenatal OP metabolite levels and the Wechsler Working Memory Index at 7 years (Bouchard et al. 2011, Rauh et al. 2011), but one study did not (Engel et al. 2011), and another found no association based on child DAP levels (Lizardi et al. 2008). In addition, one study found an inverse association with the Wechsler Perceptual Reasoning Index at age 7 years (Bouchard et al. 2011) and another detected that association among $PON1_{192}$ QQ carriers (Engel et al. 2011), whereas no significant associations were detected in the other two studies (Lizardi et al. 2008, Rauh et al. 2011). Three of four studies detected no significant associations between prenatal or child OP metabolite levels and the Wechsler Full-Scale IQ, Processing Speed, and Verbal Comprehension Scales (Engel et al. 2011, Lizardi et al. 2008, Rauh et al. 2011).

Six studies in five settings evaluated ADHD and other attention problems in preschool- and school-aged children using the Child Behavior Checklist, the NEPSY visual attention subtest, the Conners' Parental Rating Scales and Continuous Performance Test, the Hillside Behavior Rating Scale, composite ADHD indices, the Diagnostic Interview Schedule for Children IV, the Strengths and Difficulties Questionnaire, and the Behavior Assessment System for Children (Bouchard et al. 2010, Eskenazi et al. 2007, Fortenberry et al. 2014, Marks et al. 2010, Oulhote and Bouchard 2013, Rauh et al. 2006). Significant positive associations between prenatal or child OP metabolite levels and some (but not all, in the case of CHAMACOS) measures of ADHD or attention problems were detected in the CCCEH study at age 36 months (Rauh et al. 2006), in the CHAMACOS cohort at age 5 years (Marks et al. 2010), and in NHANES at ages 8–15 years (Bouchard et al. 2010), but not in the CHAMACOS cohort at ages 24 months and 3.5 years (Eskenazi et al. 2007, Marks et al. 2010), the Canadian Health Measures Survey at ages 6–11 years (Oulhote and Bouchard 2013), or the ELEMENT study at ages 6–11 years (Fortenberry et al. 2014). The consistency of results across studies is difficult to judge, due to differences in measurement instruments, analytic approaches, the timing of metabolite measurement, and the timing of neurodevelopmental assessment; the internal inconsistency of the findings in the CHAMACOS cohort also complicate interpretation. Overall, the findings for ADHD and attention problems were approximately equally balanced between positive and null.

Other behavioral problems in preschool- and school-aged children were measured in three study settings using the Child Behavior Checklist and the Strengths and Difficulties Questionnaire (Eskenazi et al. 2010, Eskenazi et al. 2007, Oulhote and Bouchard 2013, Rauh et al. 2006). The only specific behavioral outcome measured in more than one study was pervasive developmental disorder based on the Child Behavior Checklist, which was positively associated with pre- or perinatal OP metabolite levels in the CCCEH study (Rauh et al. 2006) and the CHAMACOS study (Eskenazi et al. 2007). Although scores from the Strengths and Difficulties Questionnaire have been shown to be highly correlated with those from the Child Behavior Checklist (Goodman and Scott 1999), it is unclear whether the global total difficulties scale—which was not significantly associated with child urinary DAP, DMP, or DEP metabolite levels in the Canadian Health Measures Survey (Oulhote and Bouchard 2013)—is comparable to that for pervasive developmental disorder based on the Child Behavior Checklist.

In summary, multiple studies reported a variety of associations of OP metabolites with poorer reflexes in neonates and working memory, perceptual reasoning, measures of ADHD and attention problems, and pervasive developmental disorder in school-aged children. However, this apparent consistency was detected among only three studies for neonatal reflexes and selected measures of ADHD and attention problems, and only two studies for working memory, perceptual reasoning, and pervasive developmental disorder. In addition, there were no associations across three studies for adverse neonatal behavioral outcomes, infant psychomotor development, other measures of ADHD and attention problems in school-aged children, and full-scale IQ, processing speed, and verbal comprehension in school-aged children.

When studies were closely compared according to design, exposure metric, timing of exposure measurement, age group of subjects, and neurodevelopmental test, at most only two studies were directly comparable. That is, the Mount Sinai CECS (Engel et al. 2007 2011) and CHAMACOS (Young et al. 2005, Eskenazi et al. 2007, Bouchard et al. 2011) both used a prospective cohort study design to evaluate prenatal maternal DAP, DMP, and DEP levels in association with neurodevelopmental outcomes measured using the Brazelton Neonatal Behavioral Assessment Scale, the Bayley Scales of Infant Development, and the Wechsler Intelligence Scale. Other prospective birth cohort studies used some of these neurodevelopmental tests but not the same exposure metrics (Rauh et al. 2006 2011, Lovasi et al. 2011), the same exposure metrics but different neurodevelopmental tests (Yolton et al. 2013), or different exposure and outcome measures (Forten-

berry et al. 2014). Thus, using our *a priori* requirement of three independent studies to evaluate the weight of epidemiologic evidence (stated in the "Scope of review" section), the available data are insufficient to establish consistent associations between specific OP metabolites and specific neurodevelopmental outcomes.

Temporality. Issues related to the temporality of measured OP metabolites and neurodevelopmental outcomes are similar to those discussed above in the evaluation of associations with birth outcomes, except that perinatal or early childhood exposures could plausibly be related to subsequent neurodevelopment. However, because little is known about the timing of various neurodevelopmental impairments, it is unclear whether environmental exposures in early gestation, late gestation, infancy, early childhood, or later childhood—or perhaps a combination of these—are most etiologically relevant. The paucity of knowledge about possible biological mechanisms and latency periods in neurodevelopment may justify the practice of multiple comparisons, with exposures and outcomes measured at multiple time periods being tested for associations. However, the pitfalls of this approach should be acknowledged; it is scientifically invalid to test numerous associations and choose the statistically significant ones as being the etiologically correct ones while dismissing the statistically non-significant associations.

Biological gradient. Although most studies implicitly assumed a log-linear exposure–outcome relationship between OP metabolites and neurodevelopmental outcomes, several explicitly tested for a **biological gradient** by categorizing exposures into at least three ordinal categories and assessing trends across those categories, with mixed results. Specifically, in the Mount Sinai CECS, although positive associations were detected between a continuous increase in maternal prenatal urinary DAP and DEP levels and number of abnormal reflexes in neonates, the estimated association with DAP concentrations was stronger for the second-lowest quartile than the highest (compared with the lowest), and the association with DEP concentrations was stronger for the second-highest quartile than the highest, suggesting a non-monotonic relationship (Engel et al. 2007). In the same study, the inverse association observed between maternal prenatal DAP and DMP levels and the Bayley Mental Development Index at 12 months among black or Hispanic infants appeared to be monotonic when evaluated across tertiles of metabolite levels (Engel et al. 2011). In the CHAMACOS cohort, exposure–response gradients were tested between quintiles of maternal prenatal urinary DAP, DMP, and DEP levels and >3 versus ≤ 3 abnormal reflexes in infants aged >3 days to ≤ 2 months (Young et al. 2005), as well as between quintiles of maternal prenatal urinary DAP levels and all five Wechsler Intelligence Scale measures at age 7 years (Bouchard et al. 2011). Although Wechsler scores were consistently lower in the second than the third quartile of prenatal DAP levels, significant monotonic trends were detected in all of these analyses, supporting their validity. No significant trends were detected between ordinal categories of maternal prenatal urinary MDA and TCPy levels and the Bayley Motor and Psychosocial Development Indices at 6, 12, and 24 months (Eskenazi et al. 2007).

In NHANES data, ordinally increasing categories—undetectable, below median, or above median—of urinary DMTP (the only metabolite evaluated as a categorical variable) were associated with progressively higher ORs for parent-reported prevalent ADHD, suggesting a monotonic gradient (Bouchard et al. 2010). In school-aged Mexican children, no statistically significant monotonic trends were detected across increasing tertiles of maternal prenatal urinary TCPy levels and various screening measures of ADHD (Fortenberry et al. 2014). Borderline significant trends were detected with increasing scores on the Conners' Parental ADHD index among boys ($P_{trend} = 0.06$) and increasing hit reaction time block change on the Conners' Continuous Performance Test among boys and girls combined, but evidence of non-monotonicity was detected for both outcomes among girls. Finally, analyses based on maternal prenatal urinary DAP levels categorized into quintiles clearly illustrated monotonic inverse associations with the behavior, passive tone, active tone, primary reflexes, and summary scores on the Neonatal Behavioral Neurological Assessment among Chinese neonates, supporting the results of linear regression models (Zhang et al. 2014). Overall, then, most of this subset of studies detected biological gradients that strengthen the evidence in support of exposure–response relationships between OP insecticide exposure and adverse neurodevelopmental outcomes, although some results were not consistent with monotonic trends.

Plausibility and coherence. The biological plausibility and coherence of the epidemiologic and toxicological evidence on OP insecticides in relation to neurodevelopmental outcomes was discussed above. The OP insecticide levels measured in the epidemiologic studies are far lower than would cause meaningful AChE inhibition based on animal and (limited) human toxicology data, and lower than has been established as clinically significant in animal studies. There are no known biological pathways for OP insecticides to cause the neurodevelopmental effects examined in the epidemiologic studies. Although the lack of established pathways does not mean that they do not exist, the existing evidence does not support a causal interpretation.

A consideration in the evaluation of coherence of evidence is whether observed interactions between OP metabolite levels and PON1 activity levels or genotypes suggest greater susceptibility to adverse neurodevelopmental effects of OP insecticides in individuals with lower PON1 activity levels. Three studies evaluated such interactions. In the Mount Sinai CECS, significantly stronger positive associations between maternal prenatal DAP and DMP metabolite levels and having ≥ 2 versus <2 abnormal neonatal reflexes were detected among those with lower levels of maternal plasma PON1 enzymatic activity, with an increasing exposure–response pattern in the RR across tertiles of decreasing PON1 activity (Engel et al. 2007). By contrast, in the same cohort, mixed results were obtained in analyses of interactions of maternal prenatal urinary DAP, DMP, and DEP levels with $PON1_{192}$, $PON1_{L55M}$, and $PON1_{-108C>T}$ polymorphisms and PON1 enzymatic activity (Engel et al. 2011). In particular, $PON1_{192}$ genotype interacted with prenatal DAP and DMP levels in the expected direction (i.e., with stronger adverse associations in R allele carriers) with respect to the Bayley Mental

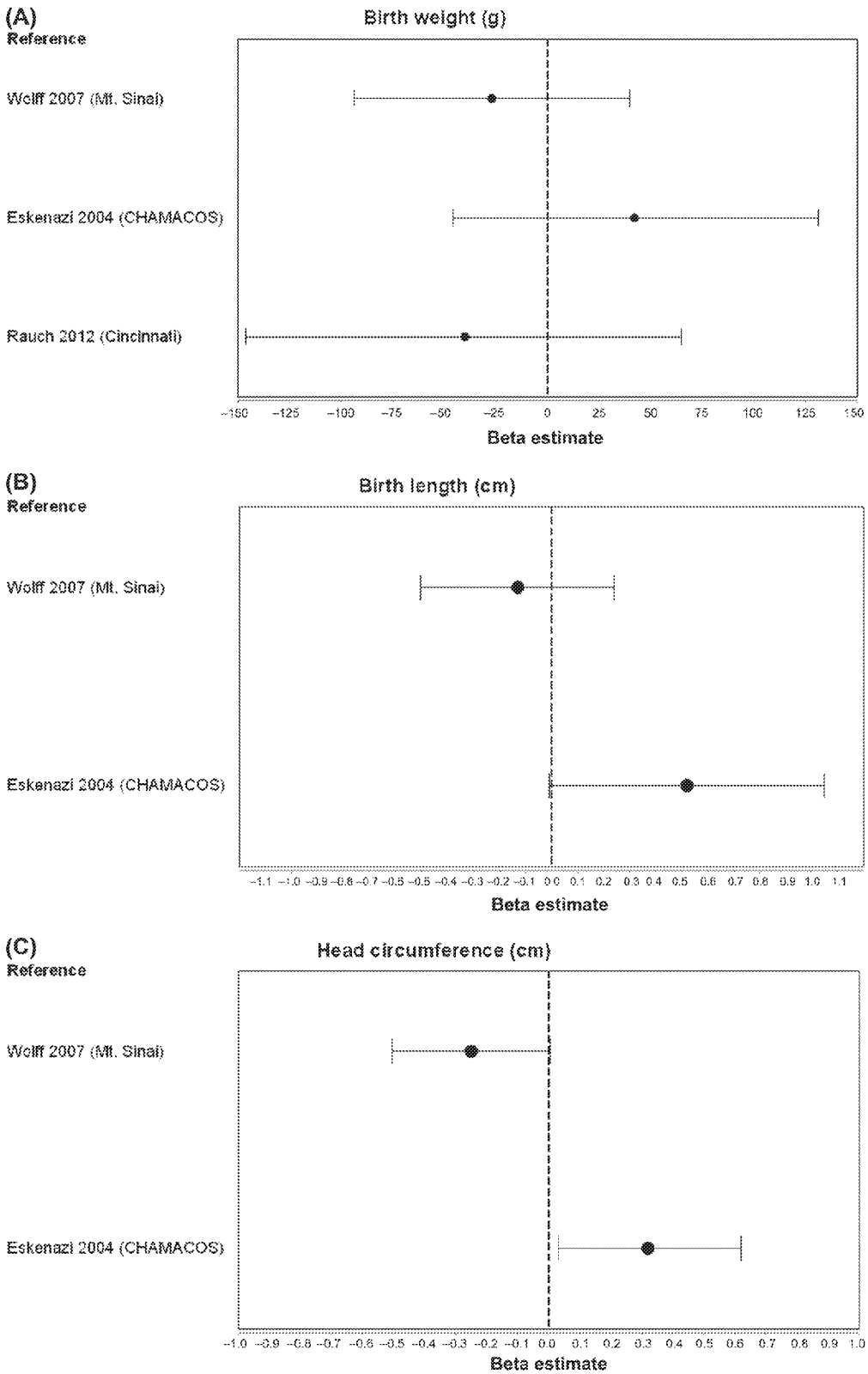


Figure 3. Estimated associations between maternal prenatal urinary DAP levels and birth outcomes. Circles indicate estimated regression coefficients (betas), with 95% CIs indicated by whiskers. Exposures are $\log_{10}$ DAP concentrations in nmol/g creatinine (Wolff et al. 2007), nmol/L (Eskenazi et al. 2004), and nmol/L, creatinine-standardized (Rauch et al. 2012). A. Associations with birth weight in grams. B. Associations with birth length in centimeters. C. Associations with head circumference in centimeters. D. Associations with ponderal index in grams per cubic centimeter. E. Associations with gestational age in weeks.

Development Index at 12 months in blacks and Hispanics, but opposite to the hypothesized direction with respect to the Wechsler Perceptual Reasoning Index at 6–9 years, and no significant interactions were found for the other PON1 genotype and enzyme measures or neurodevelopmental outcomes tested. In the CHAMACOS study, interactions were examined between maternal prenatal urinary DAP, DMP, and DEP levels and maternal and child PON1 enzyme mea-

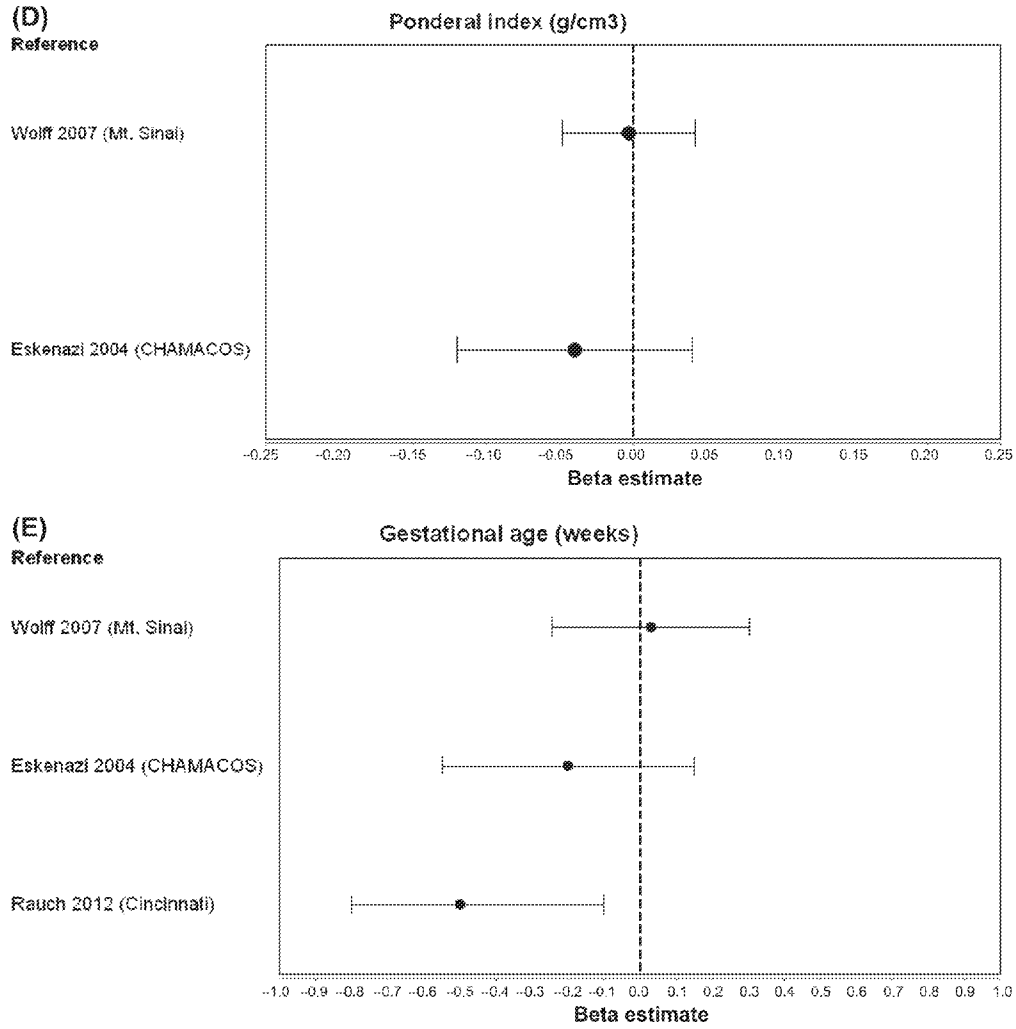


Figure 3. (*Continued*).

surements and genotypes with respect to the Bayley Mental and Psychomotor Development Indices and Child Behavior Checklist pervasive developmental disorder score (Eskenazi et al. 2010). No apparent interactions or patterns suggesting higher susceptibility with lower PON1 activity were detected. Altogether, these limited findings do not provide consistent, coherent evidence to support the hypothesis that low PON1 activity levels augment individual susceptibility to impaired neurodevelopment from OP insecticide exposure. One possible reason for the lack of consistent evidence for higher susceptibility with lower PON1 activity is that not all OP insecticides are detoxified by PON1 (Coombes et al. 2014). Additionally, Coombes et al. (2014) found that PON1 may not affect metabolism of chlorpyrifos at environmentally relevant exposures.

Specificity, experiment, and analogy. As discussed in the evaluation of the weight of evidence on OP insecticides and birth outcomes, no specificity is evident in the relationships between any particular OP insecticide and any particular neurodevelopmental outcome. Relevant experimental or quasi-experimental evidence pertaining to low-dose OP insecticide exposure and adverse neurodevelopmental outcomes in humans is lacking, and analogies to other neurotoxic or non-neurotoxic prenatal exposures do not convincingly confirm or negate a causal hypothesis.

Discussion

This paper reviews a large body of epidemiologic literature and weighs the overall evidence using the framework of the Bradford Hill guidelines. In this section, we focus on three prospective cohort studies (CECS, CHAMACOS, and HOME) that we judged to have the most informative design, and that measured maternal prenatal urinary DAP levels prior to birth or neurodevelopmental outcomes. In addition, we focused on associations with total urinary DAPs in all study subjects combined, to facilitate comparisons across studies, because DAPs were the common exposure metric. Where available, we used associations with creatinine-standardized urinary DAP levels and those that were fully adjusted for potential confounders.

As shown in Figure 3, these most informative cohort studies on balance reported no consistent significant associations between maternal prenatal urinary DAP levels and birth weight, birth length, head circumference, ponderal index, or gestational age. Figure 4 shows that, in general, there was also a lack of a significant association between maternal prenatal urinary DAP levels and most measures of neonatal neurodevelopment. Urinary DAP levels were significantly associated with increased risk of abnormal reflexes at birth in two studies (Engel et al. 2007, Young

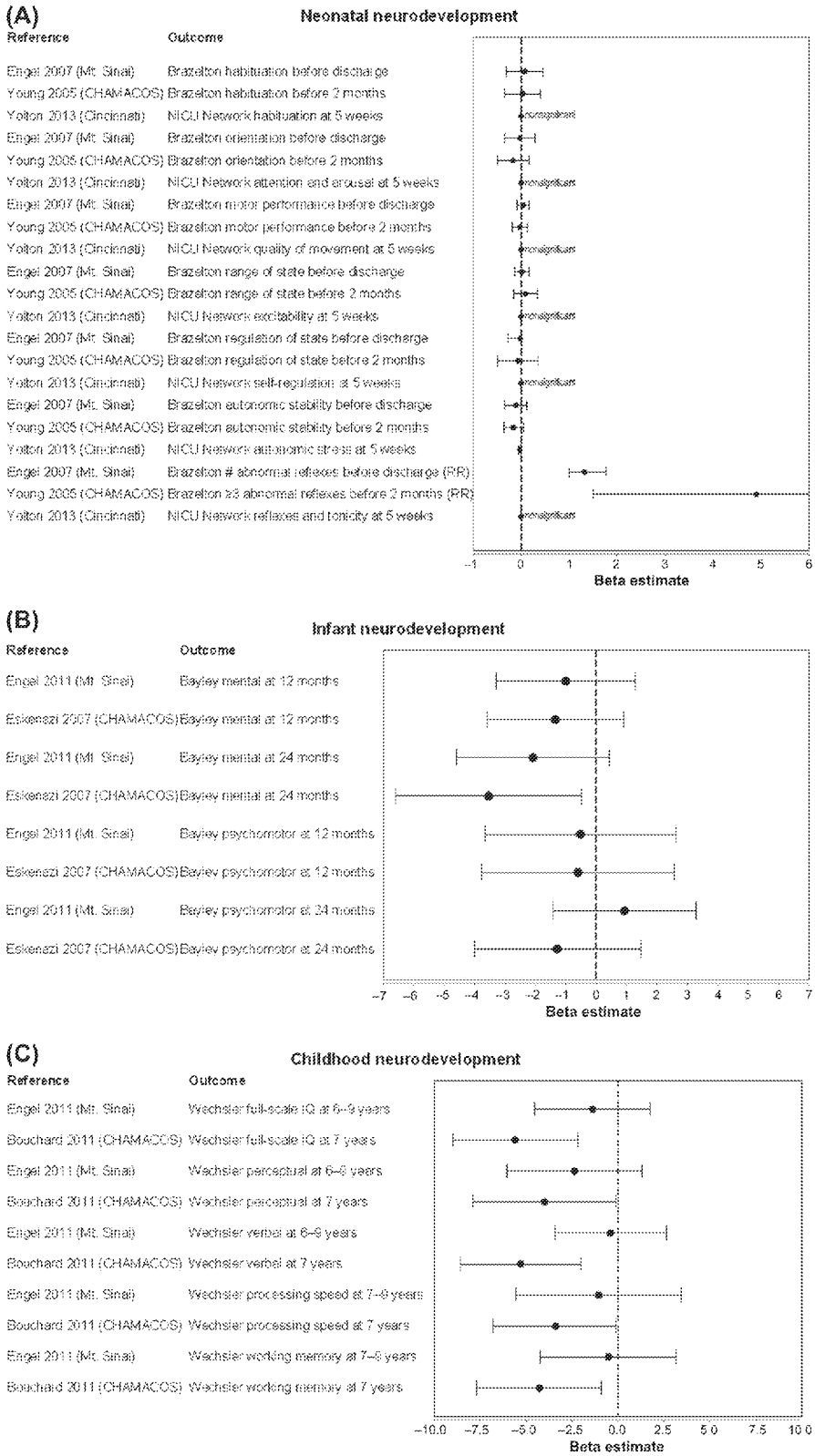


Figure 4. Estimated associations between maternal prenatal urinary DAP levels and neurodevelopmental outcomes. Circles indicate estimated regression coefficients (betas), except for associations with abnormal reflexes in neonates, where circles indicate estimated relative risks. Whiskers indicate 95% CIs. Exposures are $\log_{10}$ DAP concentrations in nmol/L (Engel et al. 2007, Engel et al. 2011, Young et al. 2005, and Bouchard et al. 2011) and in nmol/g creatinine (converted from $\log_2$) (Yolton et al. 2013). A. Associations with neurodevelopmental outcomes in neonates. Most quantitative estimates were not reported by Yolton et al. (2013), who stated that associations not shown were statistically non-significant. B. Associations with neurodevelopmental outcomes in infants. C. Associations with neurodevelopmental outcomes in children.

et al. 2005), though not in the third cohort (Yolton et al. 2013). No consistent significant associations were detected between maternal prenatal urinary DAPs and Bayley measures of neurodevelopment in infancy. Several significant associations between prenatal urinary DAPs and Wechsler measures of cognitive development in childhood

were detected in one study (Bouchard et al. 2011), but not the other (Engel et al. 2011), although point estimates in the latter study were below the null value.

Overall, these three most informative and comparable studies did not establish any consistent associations between maternal prenatal urinary DAP levels and birth or neurodevelopmental outcomes. Although results for abnormal neonatal reflexes and poorer childhood cognitive development suggested a possible association, these were not entirely consistent across studies and require independent confirmation.

In summary, associations observed between OP metabolites and birth outcomes in epidemiologic studies have been mostly weak or imprecise, inconsistent, temporally ambiguous, not clearly monotonic, not biologically plausible or coherent with toxicological evidence given the estimated degree of AChE inhibition at observed DAP concentrations, and not specific to any OP insecticide or health outcome. Associations between OP metabolites and neurodevelopmental outcomes in observed in these epidemiologic studies likewise do not unequivocally meet any of the Bradford Hill guidelines. Sir Austin Bradford Hill stated that none of these guidelines must necessarily be met to establish a causal relationship: "None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*" (Hill 1965). Some might consider the standards for causality used in this analysis as restrictive (e.g., use of Bradford Hill, criteria for agreement across three independent studies). It is possible that the use of other standards could yield different conclusions.

The inconsistencies across the studies also have to be considered in light of the lack of a biologically plausible mechanism for the adverse birth outcomes or neurodevelopmental effects evaluated in the studies. Even far less severe effects, such as mild AChE inhibition, occur at dosages that are substantially higher than the OP insecticide levels measured in the epidemiologic studies.

A common limitation of existing studies is their reliance on non-specific DAP metabolites measured in one or two urine specimens. Given the high variability of DAP metabolite levels in time-series analysis, more frequent sampling is needed to more accurately estimate exposure during pregnancy. In addition, studies with repeated serum or plasma measurements could provide further insight into the relationship of OP exposure with birth outcomes, childhood growth, and neurodevelopment. Standardization of exposures and neurodevelopmental measures would also aid comparisons across studies. In general, studies must be larger to enable statistically robust analyses of gene/environment interactions; to this end, pooling of study populations might be useful. However, efforts should be made to recognize and adjust for the expected frequency of false-positive results that arise due to multiple comparisons, especially in exploratory analyses (Wacholder et al. 2004, Glickman et al. 2014). In light of the existing limitations and inconsistency of studies, the body of epidemiologic data available at this time does not convincingly demonstrate an effect of low-level OP insecticide exposure on any adverse health outcomes in humans.

Although or goal was to include all relevant data on this topic, it is possible that some studies published in non-English journals were missed in our review. It is also possible that the current literature is subject to publication and reporting bias. It is has been shown empirically that null results in general are less likely to be reported (Dickersin and Min 1993, Dwan et al. 2008), or if reported, presented in the conclusions (Kyzas et al. 2007).

Conclusions

Recent epidemiologic studies on balance have found weak and inconsistent associations of maternal exposures to OP insecticides with birth outcome and neurodevelopmental testing results in the offspring. Perhaps the most important limitation of the extant literature is the exposure classification, which is subject to significant uncertainties due to limited sampling during pregnancy, despite the high temporal variability in exposure. In addition, the available studies cannot differentiate metabolites that form directly on and in food items and are not the result of OP insecticide exposure. Given the heterogeneity across studies in terms of overall design, types of exposure biomarkers assessed, timing of exposure measurement, birth outcomes, neurodevelopmental tests, statistical modeling approaches, and reporting of results, inter-study comparisons are challenging, and consistency of findings has not been established.

The available toxicology data show that the dosages required to cause AChE inhibition are far higher than the levels observed in the epidemiologic studies, a finding that raises further uncertainties about the biological plausibility of the epidemiologic findings. Nonetheless, the studies evaluate potential effects of major public health importance, and some of the findings, particularly poorer reflexes in neonates, ADHD/attention problems, lower cognitive scores in preschool or school-aged children, and changes in brain morphology, warrant additional study.

Acknowledgements

The authors gratefully acknowledge the review and comments from six anonymous reviewers selected by the editor. The manuscript was significantly strengthened as the result of these reviews. The authors also gratefully acknowledge Mr. Rick Nelson at Exponent for providing a careful editing of the manuscript.

Declaration of interests

The research reported in this paper was sponsored by CropLife America, which represents agrochemical companies that manufacture OP insecticides, under a contract with Exponent. RJR and MG were subcontractors to Exponent. CropLife America reviewed the paper, but the authors had ultimate authority to determine its content. The author's affiliations are listed on the cover page. RR and ETC work for Exponent, which performs risk assessment consulting for companies that produce OP insecticides. RR has served as an expert witness on behalf of defendants for cases involving OP insecticides, though not related to the topics in this paper. RJR has been the Principal Investigator on research grants and research donations to the University of Michigan from Dow Chemical Company and Dow AgroSciences. He has also been appointed by the University of Michigan as the Dow Professor of Toxicology, a professorship endowed by

the Dow Foundation. In addition, he has served as a consultant and expert witness on behalf of Dow Chemical Company and Dow AgroSciences.

References

Abu-Saad K, Fraser D. (2010). Maternal nutrition and birth outcomes. Epidemiol Rev, 32, 5–25.

Andres RL, Day MC. (2000). Perinatal complications associated with maternal tobacco use. Semin Neonatol, 5, 231–41.

Anjos T, Altmae S, Emmett P, Tiemeier H, Closa-Monasterolo R, Luque V, et al. (2013). Nutrition and neurodevelopment in children: focus on NUTRIMENTHE project. Eur J Nutr, 52, 1825–42.

Attfield KR, Hughes MD, Spengler JD, Lu C. (2014). Within- and between-child variation in repeated urinary pesticide metabolite measurements over a 1-year period. Environ Health Perspect, 122, 201–6.

Aviram M, Vaya J. (2013). Paraoxonase 1 activities, regulation, and interactions with atherosclerotic lesion. Curr Opin Lipidol, 24, 339–44.

Barr DB, Ananth CV, Yan X, Lashley S, Smulian JC, Ledoux TA, et al.. (2010). Pesticide concentrations in maternal and umbilical cord sera and their relation to birth outcomes in a population of pregnant women and newborns in New Jersey. Sci Total Environ, 408, 790–5.

Berkowitz GS, Wetmur JG, Birman-Deych E, Obel J, Lapinski RH, Godbold JH, et al. (2004). In utero pesticide exposure, maternal paraoxonase activity, and head circumference. Environ Health Perspect, 112, 388–91.

Berman T, Goldsmith R, Goen T, Spungen J, Novack L, Levine H, et al. (2013). Urinary concentrations of organophosphate pesticide metabolites in adults in Israel: demographic and dietary predictors. Environ Int, 60, 183–9.

Bliddal M, Olsen J, Stovring H, Eriksen HL, Kesmodel US, Sorensen TI, Nohr EA. (2014). Maternal pre-pregnancy BMI and intelligence quotient (IQ) in 5-year-old children: a cohort based study. PLoS One, 9, e94498.

Boon PE, Van der Voet H, Van Raaij MT, Van Klaveren JD. (2008). Cumulative risk assessment of the exposure to organophosphorus and carbamate insecticides in the Dutch diet. Food Chem Toxicol, 46, 3090–8.

Bouchard M, Gosselin NH, Brunet RC, Samuel O, Dumoulin MJ, Carrier G. (2003). A toxicokinetic model of malathion and its metabolites as a tool to assess human exposure and risk through measurements of urinary biomarkers. Toxicol Sci, 73, 182–94.

Bouchard MF, Bellinger DC, Wright RO, Weisskopf MG. (2010). Attention-deficit/hyperactivity disorder and urinary metabolites of organophosphate pesticides. Pediatrics, 125, e1270–7.

Bouchard MF, Chevrier J, Harley KG, Kogut K, Vedar M, Calderon N, et al. (2011). Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. Environ Health Perspect, 119, 1189–95.

Bradman A, Eskenazi B, Barr DB, Bravo R, Castorina R, Chevrier J, et al. (2005). Organophosphate urinary metabolite levels during pregnancy and after delivery in women living in an agricultural community. Environ Health Perspect, 113, 1802–7.

Bradman A, Kogut K, Eisen EA, Jewell NP, Quiros-Alcala L, Castorina R, et al. (2013). Variability of organophosphorous pesticide metabolite levels in spot and 24-hr urine samples collected from young children during 1 week. Environ Health Perspect, 121, 118–24.

Burkhalter TM, Hillman CH. (2011). A narrative review of physical activity, nutrition, and obesity to cognition and scholastic performance across the human lifespan. Adv Nutr, 2, 201S–6S.

Burns CJ, McIntosh LJ, Mink PJ, Jurek AM, Li AA. (2013). Pesticide exposure and neurodevelopmental outcomes: review of the epidemiologic and animal studies. J Toxicol Environ Health B Crit Rev, 16, 127–283.

Cantor KP, Blair A, Everett G, Gibson R, Burmeister LF, Brown LM, et al. (1992). Pesticides and other agricultural risk factors for non-Hodgkin's lymphoma among men in Iowa and Minnesota. Cancer Res, 52, 2447–55.

Carr RL, Adams AL, Kepler DR, Ward AB, Ross MK. (2013). Induction of endocannabinoid levels in juvenile rat brain following developmental chlorpyrifos exposure. Toxicol Sci, 135, 193–201.

Centers for Disease Control and Prevention (CDC). (2014). National Health and Nutrition Examination Survey. CDC Website. http://www.cdc.gov/nchs/nhanes.htm.

Chen L, Zhao T, Pan C, Ross J, Ginevan M, Vega H, Krieger R. (2013). Absorption and excretion of organophosphorous insecticide biomarkers of malathion in the rat: implications for overestimation bias and exposure misclassification from environmental biomonitoring. Regul Toxicol Pharmacol, 65, 287–93.

Chen L, Zhao T, Pan C, Ross JH, Krieger RI. (2012). Preformed biomarkers including dialkylphosphates (DAPs) in produce may confound biomonitoring in pesticide exposure and risk assessment. J Agric Food Chem, 60, 9342–51.

Claeys WL, De Voghel S, Schmit JF, Vromman V, Pussemier L. (2008). Exposure assessment of the Belgian population to pesticide residues through fruit and vegetable consumption. Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 25, 851–63.

Coombes RH, Meek EC, Dail MB, Chambers HW, Chambers JE. (2014). Human paraoxanase 1 hydrolysis of nanomolar chlorpyrifos-oxon concentrations is unaffected by phenotype or Q192R genotype. Toxicol Lett, 230, 57–61.

Curl CL, Fenske RA, Elgethun K. (2003). Organophosphorous pesticide exposure of urban and suburban preschool children with organic and conventional diets. Environ Health Perspect, 111, 377–82.

Dickersin K, Min YI. (1993). Publication bias: the problem that won't go away. Ann NY Acad Sci, 703, 135–46; discussion 146–138.

Duggan A, Charnley G, Chen W, Chukwudebe A, Hawk R, Krieger RL, et al. 2003. Dialkylphosphate biomonitoring data: assessing cumulative exposure to organophosphate pesticides. Regul Toxicol Pharmacol, 37, 382–95.

Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One, 3, e3081.

Gonzales ET, Bauer SB. (1999). Pediatric urology practice. Philadelphia, PA: Lippincott Williams , & Wilkins.

Eaton DL, Daroff RB, Autrup H, Bridges J, Buffler P, Costa LG, et al. (2008). Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. Crit Rev Toxicol, 38, 1–125.

Eddleston M, Buckley NA, Eyer P, Dawson AH. (2008). Management of acute organophosphorus pesticide poisoning. Lancet, 371, 597–607.

Engel SM, Berkowitz GS, Barr DB, Teitelbaum SL, Siskind J, Meisel SJ, et al. (2007). Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton Neonatal Behavioral Assessment Scale in a multiethnic pregnancy cohort. Am J Epidemiol, 165, 1397–404.

Engel SM, Wetmur J, Chen J, Zhu C, Barr DB, Canfield RL, Wolff MS. (2011). Prenatal exposure to organophosphates, paraoxonase 1, and cognitive development in childhood. Environ Health Perspect, 119, 1182–8.

Eskenazi B, Harley K, Bradman A, Weltzien E, Jewell NP, Barr DB, et al. (2004). Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. Environ Health Perspect, 112, 1116–24.

Eskenazi B, Huen K, Marks A, Harley KG, Bradman A, Barr DB, Holland N. (2010). PON1 and neurodevelopment in children from the CHAMACOS study exposed to organophosphate pesticides in utero. Environ Health Perspect, 118, 1775–81.

Eskenazi B, Marks AR, Bradman A, Harley K, Barr DB, Johnson C, et al. (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. Environ Health Perspect, 115, 792–8.

Ferioli A, Maroni M. (2011). Biological monitoring. In: La Ferla F, Lauwerys RR, eds. Encyclopedia of Occupational Health and Safety. Geneva, Switzerland: International Labor Organization.

Fortenberry GZ, Meeker JD, Sanchez BN, Barr DB, Panuwet P, Bellinger D, et al. (2014). Urinary 3,5,6-trichloro-2-pyridinol (TCPY) in pregnant women from Mexico City: distribution, temporal variability, and relationship with child attention and hyperactivity. Int J Hyg Environ Health, 217, 405–12.

Garfitt SJ, Jones K, Mason HJ, Cocker J. (2002). Exposure to the organophosphate diazinon: data from a human volunteer study with oral and dermal doses. Toxicol Lett, 134, 105–13.

Giles D, Dickson J. (2000). A randomized double blind ascending single oral dose study with malathion to determine the no effect level on plasma and RBC cholinesterase activity. Tranent, Scotland: Inveresk Research,

Glickman ME, Rao SR, Schultz MR. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. J Clin Epidemiol, 67, 850–7.

Gonzales ET, Bauer SB. (1999). Pediatric urology practice. Philadelphia, PA: Lippincott Williams & Wilkins.

Goodman R, Scott S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: is small beautiful? J Abnorm Child Psychol, 27, 17–24.

Gordis L. (2013). Epidemiology. 5th ed. Philadelphia: Elsevier.

Griffith W, Curl CL, Fenske RA, Lu CA, Vigoren EM, Faustman EM. (2011). Organophosphate pesticide metabolite levels in pre-school children in an agricultural community: within- and between-child variability in a longitudinal study. Environ Res, 111, 751–6.

Guodong D, Pei W, Ying T, Jun Z, Yu G, Xiaojin W, et al. (2012). Organophosphate pesticide exposure and neurodevelopment in young Shanghai children. Environ Sci Technol, 46, 2911–7.

Harley KG, Huen K, Aguilar Schall R, Holland NT, Bradman A, Barr DB, Eskenazi B. (2011). Association of organophosphate pesticide exposure and paraoxonase with birth outcome in Mexican-American women. PLoS One, 6, e23923.

Hill AB. (1965). The environment and disease: association or causation?. Proc R Soc Med, 58, 295–300.

Horton MK, Kahn LG, Perera F, Barr DB, Rauh V. (2012). Does the home environment and the sex of the child modify the adverse effects of prenatal exposure to chlorpyrifos on child working memory? Neurotoxicol Teratol, 34, 534–41.

Jensen AF, Petersen A, Granby K. (2003). Cumulative risk assessment of the intake of organophosphorus and carbamate pesticides in the Danish diet. Food Addit Contam, 20, 776–85.

Jensen BH, Petersen A, Christensen T. (2009). Probabilistic assessment of the cumulative dietary acute exposure of the population of Denmark to organophosphorus and carbamate pesticides. Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 26, 1038–48.

Jurek AM, Greenland S, Maldonado G. (2008). How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? Int J Epidemiol, 37, 382–5.

Jurek AM, Greenland S, Maldonado G, Church TR. (2005). Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol, 34, 680–7.

Krasowski MD, Mcgehee DS, Moss J. (1997). Natural inhibitors of cholinesterases: implications for adverse drug reactions. Can J Anaesth, 44, 525–34.

Krieger RI, Chen L, Ginevan M, Watkins D, Cochran RC, Driver JH, Ross JH. 2012. Implications of estimates of residential organophosphate exposure from dialkylphosphates (DAPs) and their relevance to risk. Regul Toxicol Pharmacol, 64, 263–66.

Kyzas PA, Denaxa-Kyza D, Ioannidis JP. (2007). Almost all articles on cancer prognostic markers report statistically significant results. Europ J Cancer, 43, 2559–2579.

Lefkowitz LJ, Kupina JM, Hirth NL, Henry RM, Noland GY, Barbee JYJ, et al. (2007). Intraindividual stability of human erythrocyte cholinesterase activity. Clin Chem, 53, 1358–1363.

Lester BM, Tronick EZ, Brazelton TB. (2004). The Neonatal Intensive Care Unit Network Neurobehavioral Scale procedures. Pediatrics, 113, 641–67.

Li B, Sedlacek M, Manoharan I, Boopathy R, Duysen EG, Masson P, Lockridge O (2005). Butyrylcholinesterase, paraoxonase, and albumin esterase, but not carboxylesterase, are present in human plasma. Biochem Pharmacol, 70, 1673–1684.

Li AA, Lowe KA, McIntosh LJ, Mink PJ. (2012). Evaluation of epidemiology and animal data for risk assessment: chlorpyrifos developmental neurobehavioral outcomes. J Toxicol Environ Health B Crit Rev, 15, 109–84.

Lizardi PS, O'Rourke MK, Morris RJ. (2008). The effects of organophosphate pesticide exposure on Hispanic children's cognitive and behavioral functioning. J Pediatr Psychol, 33, 91–101.

Lovasi GS, Quinn JW, Rauh VA, Perera FP, Andrews HF, Garfinkel R, et al. (2011). Chlorpyrifos exposure and urban residential environment characteristics as determinants of early childhood neurodevelopment. Am J Public Health, 101, 63–70.

Lu C, Barr DB, Pearson MA, Waller LA. (2008). Dietary intake and its contribution to longitudinal organophosphorus pesticide exposure in urban/suburban children. Environ Health Perspect, 116, 537–42.

Macharia M, Kengne AP, Blackhurst DM, Erasmus RT, Matsha TE. (2014). The impact of chronic untreated hyperglycaemia on the long-term stability of paraoxonase 1 (PON1) and antioxidant status in human sera. J Clin Pathol, 67, 55–9.

Medina JL, Reinicke K, Simpfendörfer R, Roa A, Oliveros H, Bardisa L, Rudolph MI. (1993). Characterization and distribution of cholinesterase activity in mouse uterine horns: changes in estrous cycle. Comp Biochem Phys C, 106, 473–78.

Marks AR, Harley K, Bradman A, Kogut K, Barr DB, Johnson C, et al. (2010). Organophosphate pesticide exposure and attention in young Mexican-American children: the CHAMACOS study. Environ Health Perspect, 118, 1768–74.

Marshall NE, Spong CY. (2012). Obesity, pregnancy complications, and birth outcomes. Semin Reprod Med, 30, 465–71.

Mason JB, Saldanha LS, Ramakrishnan U, Lowe A, Noznesky EA, Girard AW, et al. (2012). Opportunities for improving maternal nutrition and birth outcomes: synthesis of country experiences. Food Nutr Bull, 33, S104–37.

McDaniel CY, Dail MB, Wills RW, Chambers HW, Chambers JE. (2014) Paraoxonase 1 polymorphisms within a Mississippi USA population as possible biomarkers of enzyme activities associated with disease susceptibility. Biochem Genet, 52, 509–23.

Mileson BE, Chambers JE, Chen WL, Dettbarn W, Ehrich M, Eldefrawi AT, et al. (1998) Common mechanism of toxicity: a case study of organophosphorus pesticides. Toxicol Sci, 41, 8–20.

Millichap JG, Yee MM. (2012). The diet factor in attention-deficit/hyperactivity disorder. Pediatrics, 129, 330–7.

Mink PJ, Kimmel CA, Li AA. (2012). Potential effects of chlorpyrifos on fetal growth outcomes: implications for risk assessment. J Toxicol Env Heal B, 15, 281–316.

Morgan MK, Sheldon LS, Jones PA, Croghan CW, Chuang JC, Wilson NK. (2011). The reliability of using urinary biomarkers to estimate children's exposures to chlorpyrifos and diazinon. J Expo Sci Environ Epidemiol, 21, 280–90.

Neggers YH, Goldenberg RL, Ramey SL, Cliver SP. (2003). Maternal prepregnancy body mass index and psychomotor development in children. Acta Obstet Gynecol Scand, 82, 235–40.

Nougadere A, Sirot V, Kadar A, Fastier A, Truchot E, Vergnet C, et al. (2012). Total diet study on pesticide residues in France: levels in food as consumed and chronic dietary risk to consumers. Environ Int, 45, 135–50.

Oulhote Y, Bouchard MF. (2013). Urinary metabolites of organophosphate and pyrethroid pesticides and behavioral problems in Canadian children. Environ Health Perspect, 121, 1378–84.

Ozarowski M, Mikolajczak PL, Bogacz A, Gryszczynska A, Kujawska M, Jodynis-Liebert J, et al. (2013). Rosmarinus officinalis L. leaf extract improves memory impairment and affects acetylcholinesterase and butyrylcholinesterase activities in rat brain. Fitoterapia, 91, 261–71.

Perera FP, Rauh V, Tsai WY, Kinney P, Camann D, Barr D, et al. (2003). Effects of transplacental exposure to environmental pollutants on birth outcomes in a multiethnic population. Environ Health Perspect, 111, 201–5.

Quiros-Alcala L, Alkon AD, Boyce WT, Lippert S, Davis NV, Bradman A, et al. (2011). Maternal prenatal and child organophosphate pesticide exposures and children's autonomic function. Neurotoxicology, 32, 646–55.

Rauch SA, Braun JM, Barr DB, Calafat AM, Khoury J, Montesano AM, et al. (2012). Associations of prenatal exposure to organophosphate pesticide metabolites with gestational age and birth weight. Environ Health Perspect, 120, 1055–60.

Rauh V, Arunajadai S, Horton M, Perera F, Hoepner L, Barr DB, Whyatt R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ Health Perspect, 119, 1196–201.

Rauh VA, Garfinkel R, Perera FP, Andrews HF, Hoepner L, Barr DB, et al. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. Pediatrics, 118, e1845–59.

Rauh VA, Perera FP, Horton MK, Whyatt RM, Bansal R, Hao X, et al. (2012). Brain anomalies in children exposed prenatally to a common organophosphate pesticide. Proc Natl Acad Sci U S A, 109, 7871–6.

Reiss R. (2012). Benchmark dose modeling for an omethoate acute comparative cholinesterase study in neonatal and adult rats. Prepared for Cheminova, Inc. Exponent report number VA10532.000 B0T0 0212 RR22.,

Rice D, Barone S Jr. (2000). Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. Environ Health Perspect, 108, 511–33.

Rothman KJ, Greenland S, Lash TL. (2012).Modern epidemiology. 3rd ed. (revised). Philadelphia, PA: Lippincott Williams & Wilkins

Sandjaja, Poh BK, Rojroonwasinkul N, Le Nyugen BK, Budiman B, Ng LO, et al. (2013). Relationship between anthropometric indicators and cognitive performance in Southeast Asian school-aged children. Br J Nutr, 110, S57–64.

Schrader C, Rimbach G. (2011). Determinants of paraoxonase 1 status: genes, drugs and nutrition. Curr Med Chem, 18, 5624–43.

Sexton K, Ryan AD. (2012) Using exposure biomarkers in children to compare between-child and within-child variance and calculate correlations among siblings for multiple environmental chemicals. J Expo Sci Environ Epidemiol ,22, 16–23.

Slotkin TA, Seidler FJ. (2007). Comparative developmental neurotoxicity of organophosphates in vivo: transcriptional responses of pathways for brain cell development, cell signaling, cytotoxicity and neurotransmitter systems. Brain Res Bull, 72, 232–74.

Smithers LG, Lynch JW, Yang S, Dahhou M, Kramer MS. (2013). Impact of neonatal growth on IQ and behavior at early school age. Pediatrics, 132, e53–60.

Solecki R. (2002). Pesticide residues in food - 2002 - joint FAO/WHO meeting on pesticide residues - acephate. IPCS International Program on Chemical Safety [Online] Available At: http://www.Inchem.Org/Documents/Jmpr/Jmpmono/2002pr02.Htm. Accessed on 14 July 2014.

Sorahan T, Gilthorpe MS. (1994). Non-differential misclassification of exposure always leads to an underestimate of risk: an incorrect conclusion. Occup Environ Med, 51, 839–40.

Strömberg U, Björk J, Broberg K, Mertens F, Vineis P. (2008). Selection of influential genetic markers among a large number of candidates based on effect estimation rather than hypothesis testing: an approach for genome-wide association studies. Epidemiology, 19, 302–8.

Sudakin DL, Stone DL. (2011). Dialkyl phosphates as biomarkers of organophosphates: the current divide between epidemiology and clinical toxicology. Clin Toxicol, 49, 771–81.

Timchalk C, Nolan RJ, Mendrala AL, Dittenber DA, Brzak KA, Mattsson JL. (2002). A Physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate insecticide chlorpyrifos in rats and humans. Toxicol Sci, 66, 34–53.

U.S. Department of Agriculture (USDA). (2014). Pesticide data program. annual summary, calendar year 2012. In: Agriculture USDO (ed.). Washington, DC: U.S. Department of Agriculture.

U.S. Environmental Protection Agency (EPA). (2009a). Dimethoate: human health assessment scoping document in support of registration review. USEPA Office of Pesticide Program.

U.S Environmental Protection Agency (EPA). 2009b. Registration eligibility decision (RED) for malathion. USEPA Office of Pesticide Program,

U.S. Environmental Protection Agency (EPA). (2005). Dimethoate and omethoate: comparative toxicity and determination of toxicity adjustment factors. (addendum to HED nos. 0050651 and 0050901). USEPA Ofice of Pesticide Programs,

U.S. Environmental Protection Agency (EPA). (2011). Chlorpyrifos: preliminary human health risk assessment for registration review. USEPA Ofice of Pesticide Programs (ed.),

US Environmental Protection Agency (EPA). (2002). Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity. USEPA Ofice of Pesticide Programs,

US Environmental Protection Agency (EPA). (2000). Office of pesticide programs science policy on the use of data on cholinesterase inhibition for risk assessments of organophosphorous and carbamate pesticides. USEPA Office of Pesticide Programs.

Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. (2004). Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst, 96, 434–42.

Wacholder S, Hartge P, Lubin JH, Dosemeci M. (1995). Non-differential misclassification and bias towards the null: a clarification. Occup Environ Med, 52, 557–8.

Wakefield J. (2007). A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet, 81, 208–27.

Wang P, Tian Y, Wang XJ, Gao Y, Shi R, Wang GQ, et al. (2012). Organophosphate pesticide exposure and perinatal outcomes in Shanghai, China. Environ Int, 42, 100–4.

Weitkunat R, Kaelin E, Vuillaume G, Kallischnigg G. (2010). Effectiveness of strategies to increase the validity of findings from association studies: size vs. replication. BMC Med Res Methodol, 10, 47.

World Health Organization (WHO). (1996). Biological monitoring of chemical exposure in the workplace. guidelines. volume 1. WHO/HPR/OCH 96.1. Geneva.

Whyatt RM, Camann D, Perera FP, Rauh VA, Tang D, Kinney PL, et al. (2005). Biomarkers in assessing residential insecticide exposures during pregnancy and effects on fetal growth. Toxicol Appl Pharmacol, 206, 246–54.

Whyatt RM, Camann DE, Kinney PL, Reyes A, Ramirez J, Dietrich J, et al. (2002). Residential pesticide use during pregnancy among a cohort of urban minority women. Environ Health Perspect, 110, 507–14.

Whyatt RM, Rauh V, Barr DB, Camann DE, Andrews HF, Garfinkel R, et al. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. Environ Health Perspect, 112, 1125–32.

Wickerham EL, Lozoff B, Shao J, Kaciroti N, Xia Y, Meeker JD. (2012). Reduced birth weight in relation to pesticide mixtures detected in cord blood of full-term infants. Environ Int, 47, 80–5.

Wolff MS, Engel S, Berkowitz G, Teitelbaum S, Siskind J, Barr DB, Wetmur J. (2007). Prenatal pesticide and PCB exposures and birth outcomes. Pediatr Res, 61, 243–50.

Yang D, Lauridsen H, Buels K, Chi LH, La Du J, Brunn DA, et al. (2011). Chlorpyrifos-oxon disrupts zebrafish axonal growth and motor behavior. Toxicol Sci, 121, 146–59.

Ye X, Pierik FH, Hauser R, Duty S, Angerer J, Park MM, et al. (2008). Urinary metabolite concentrations of organophosphorous pesticides, bisphenol A, and phthalates among pregnant women in Rotterdam, the Netherlands: the Generation R study. Environ Res, 108, 260–7.

Yolton K, Xu Y, Sucharew H, Succop P, Altaye M, Popelar A, et al. (2013). Impact of low-level gestational exposure to organophosphate pesticides on neurobehavior in early infancy: a prospective study. Environ Health, 12, 79.

Young JG, Eskenazi B, Gladstone EA, Bradman A, Pedersen L, Johnson C, et al. (2005). Association between in utero organophosphate pesticide exposure and abnormal reflexes in neonates. Neurotoxicology, 26, 199–209.

Zhang X, Drive JH, LI Y, Ross JH, Krieger RI. (2008). Dialkylphosphates (DAPs) in fruits and vegetables may confound biomonitoring in organophosphorus insecticide exposure and risk assessment. J Agric Food Chem, 56, 10638–45.

Zhang Y, Han S, Liang D, Shi X, Wang F, Liu W, et al. (2014). Prenatal exposure to organophosphate pesticides and neurobehavioral development of neonates: a birth cohort study in Shenyang, China. PLoS One, 9, e88491.

Message

From: Janet Collins [jcollins@croplifeamerica.org]
Sent: 4/13/2018 7:03:45 PM
To: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]; Keigwin, Richard [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=151baabb6a2246a3a312f12a706c0a05-Richard P Keigwin Jr]
CC: csmith@gowanco.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=85a93dee627e495997f325593ed303eb-csmith@gowa]
Subject: RE:

You too; thank you. Can't wait to run down to see the blossoms again this afternoon.

Janet

Ex. 6

From: Beck, Nancy [mailto:Beck.Nancy@epa.gov]
Sent: Friday, April 13, 2018 2:53 PM
To: Janet Collins <jcollins@croplifeamerica.org>; Keigwin, Richard <Keigwin.Richard@epa.gov>
Cc: csmith@gowanco.com
Subject: RE:

Thanks for following up Janet.

Enjoy the beautiful warm and sunny weekend!
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
beck.nancy@epa.gov

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Friday, April 13, 2018 2:13 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>; Keigwin, Richard <Keigwin.Richard@epa.gov>
Cc: csmith@gowanco.com
Subject:

Dear Nancy and Rick- thank you for the time you dedicated to meeting with us on Wednesday morning.

During the meeting, we discussed the EPA consideration of the exposure information from the CHAMACOS study. We discussed that CHAMACOS did not report chlorpyrifos but did report on the oxons of chlorpyrifos. Attached please find a paper published in 2012 wherein you will note the authors statement (see last sentence in abstract) that oxons would not be in the peripheral tissues- thus, would not be present in the brain- brain function would not then be affected by oxons in the blood samples.

We welcome the opportunity to discuss this further, and likely will address that specific point when we provide the final study report that we have conducted to plot the data from the Columbia University study.

Thank you again.

My best

Janet E Collins, Ph.D., R.D.
Executive Vice President, Science and Regulatory Affairs
CropLife America
1156 15th Street, NW; Suite 400
Washington DC 20001

Ex. 6

Message

From: Cindy Smith [csmith@gowanco.com]
Sent: 4/11/2018 9:07:33 PM
To: Keller, Kaitlin [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=d7a6b15adfd745c6ada1c121dec27ac4-Keller, Kai]; janet collins [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usera98e8fe5]
CC: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
Subject: RE: Update Meeting Materials

Kaitlin – wow thanks very much for getting this to us so quickly and for your time today.

From: Keller, Kaitlin <keller.kaitlin@epa.gov>
Sent: Wednesday, April 11, 2018 1:54 PM
To: janet collins <jcollins@croplifeamerica.org>
Cc: Beck, Nancy <Beck.Nancy@epa.gov>; Cindy Smith <csmith@gowanco.com>
Subject: RE: Update Meeting Materials

Hi Janet,

As follow-up to the meeting today, I have attached Appendix 6 pulled from the 2014 Chlorpyrifos Human Health Risk Assessment. Section III.B. discusses co-exposure to other environmental contaminants. Also, Russell Carr (out of Mississippi, not North Carolina, but still in the south) is the researcher studying mechanistic effects, specifically related to endocannabinoids.

Thanks,
Kaitlin

Kaitlin Keller, Special Assistant
Office of Chemical Safety and Pollution Prevention
U.S. Environmental Protection Agency
(202) 564-7098

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Monday, April 09, 2018 11:39 AM
To: Keller, Kaitlin <keller.kaitlin@epa.gov>
Cc: Beck, Nancy <Beck.Nancy@epa.gov>; csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Bolen, Derrick <bolen.derrick@epa.gov>
Subject: Re: Update Meeting Materials

Thank you Kaitlin. We can make that work.

On Apr 9, 2018, at 11:15 AM, Keller, Kaitlin <keller.kaitlin@epa.gov> wrote:

> Hi Janet,
>
> Apologies for any confusion. We can do an hour here Wednesday, can you come in 9-10am?
>
> Thank you,
> Kaitlin

Kaitlin Keller, Special Assistant
Office of Chemical Safety and Pollution Prevention
U.S. Environmental Protection Agency
(202) 564-7098

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Monday, April 09, 2018 8:13 AM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Cc: csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Keller, Kaitlin <keller.kaitlin@epa.gov>; Bolen, Derrick <bolen.derrick@epa.gov>
Subject: Re: Update Meeting Materials

Thanks Nancy- my understanding is that we have one hour.

Please confirm specifically

On Apr 9, 2018, at 7:34 AM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:

> Thanks Janet.
> Should I consider the below information the agenda for our meeting. My understanding is that your group is coming in for 30 minutes on Wednesday.
>
> Thanks,
> Nancy
>
> ---
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> **Ex. 6**
> beck.nancy@epa.gov
>
> **From:** Janet Collins [mailto:jcollins@croplifeamerica.org]
> **Sent:** Monday, April 9, 2018 7:08 AM
> **To:** Beck, Nancy <Beck.Nancy@epa.gov>
> **Cc:** csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>
> **Subject:** Update Meeting Materials
> **Importance:** High
>
> Nancy- again we really appreciate you joining our Strategic Oversight Council discussion on January 25th. You may recall that we had several items we committed to follow up on for you.
>
> 1. You raised the concept of a 3rd party review of the epidemiological data that is the basis for EPA's HED Memorandum that reapplies an FQPA 10x to all organophosphate risk assessments. We wanted to highlight for you that some 3rd party reviews of those data have been conducted. I have highlighted the summaries of the following papers and provided them in their entirety if you want to review them:
>   a. Debbie Edwards Paper
>   b. Rick Reiss/Michael Goodman Paper
>   c. Gradient Paper (2015)

2. We pointed out that EPA had completed risk assessments for some organophosphates after the epidemiological data were available to them where no FQPA 10x was applied. Here are the specific examples. Here are the specific examples of organophosphates which EPA removed the FQPA 10x and did not reapply until the HED memo issued in 2015:
    a. Bensulide – EPA scoping document in 2008 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data
    b. Phosmet – EPA scoping document in 2009 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data

    Additionally, I also include a letter written in 2013 by Steve Bradbury (then Director of the Office of Pesticide Programs) regarding the use of these same epidemiology data in risk assessments. It is our contention that for EPA to reapply the FQPA 10x to a compound from which the Agency had removed it, they must have reliable and available data. The researchers have not provided the data to the Agency for the epidemiological studies that are the basis for EPA reapplying the FQPA 10x.

3. Ongoing mechanistic data. You mentioned ongoing research–possibly at ORD – to determine if there is some other mode of action occurring at doses lower than those that inhibit cholinesterase that may cause neurodevelopmental effects. Can you please provide more information about what is being done? We would just like to point that previous discussions on this topic often focused on brain rather than RBC and the focus really needs to be on RBC.

Thanks again for agreeing to meet with us on Wednesday- we appreciate it.

Janet

Ex. 6

Message

From: Janet Collins [jcollins@croplifeamerica.org]
Sent: 4/11/2018 8:57:10 PM
To: Keller, Kaitlin [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=d7a6b15adfd745c6ada1c121dec27ac4-Keller, Kai]
CC: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]; csmith@gowanco.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=85a93dee627e495997f325593ed303eb-csmith@gowa]
Subject: Re: Update Meeting Materials

Thanks so much Kaitlin- most appreciated.

And thank you for the meeting.

On Apr 11, 2018, at 4:54 PM, Keller, Kaitlin <keller.kaitlin@epa.gov> wrote:

> Hi Janet,
>
> As follow-up to the meeting today, I have attached Appendix 6 pulled from the 2014 Chlorpyrifos Human Health Risk Assessment. Section III.B. discusses co-exposure to other environmental contaminants. Also, Russell Carr (out of Mississippi, not North Carolina, but still in the south) is the researcher studying mechanistic effects, specifically related to endocannabinoids.
>
> Thanks,
> Kaitlin
>
> Kaitlin Keller, Special Assistant
> Office of Chemical Safety and Pollution Prevention
> U.S. Environmental Protection Agency
> (202) 564-7098
>
> **From:** Janet Collins [mailto:jcollins@croplifeamerica.org]
> **Sent:** Monday, April 09, 2018 11:39 AM
> **To:** Keller, Kaitlin <keller.kaitlin@epa.gov>
> **Cc:** Beck, Nancy <Beck.Nancy@epa.gov>; csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Bolen, Derrick <bolen.derrick@epa.gov>
> **Subject:** Re: Update Meeting Materials
>
> Thank you Kaitlin. We can make that work.
>
> On Apr 9, 2018, at 11:15 AM, Keller, Kaitlin <keller.kaitlin@epa.gov> wrote:
>
> > Hi Janet,
> >
> > Apologies for any confusion. We can do an hour here Wednesday, can you come in 9-10am?
> >
> > Thank you,
> > Kaitlin

Kaitlin Keller, Special Assistant
Office of Chemical Safety and Pollution Prevention
U.S. Environmental Protection Agency
(202) 564-7098

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Monday, April 09, 2018 8:13 AM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Cc: csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Keller, Kaitlin <keller.kaitlin@epa.gov>; Bolen, Derrick <bolen.derrick@epa.gov>
Subject: Re: Update Meeting Materials

Thanks Nancy- my understanding is that we have one hour.

Please confirm specifically

On Apr 9, 2018, at 7:34 AM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:

> Thanks Janet.
> Should I consider the below information the agenda for our meeting. My understanding is that your group is coming in for 30 minutes on Wednesday.
>
> Thanks,
> Nancy
>
> _____
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> **Ex. 6**
> beck.nancy@epa.gov
>
> **From:** Janet Collins [mailto:jcollins@croplifeamerica.org]
> **Sent:** Monday, April 9, 2018 7:08 AM
> **To:** Beck, Nancy <Beck.Nancy@epa.gov>
> **Cc:** csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>
> **Subject:** Update Meeting Materials
> **Importance:** High
>
> Nancy- again we really appreciate you joining our Strategic Oversight Council discussion on January 25th. You may recall that we had several items we committed to follow up on for you.
>
> 1. You raised the concept of a 3rd party review of the epidemiological data that is the basis for EPA's HED Memorandum that reapplies an FQPA 10x to all organophosphate risk assessments. We wanted to highlight for you that some 3rd party reviews of those data have been conducted. I have highlighted the summaries of the following

papers and provided them in their entirety if you want to review them:

a. Debbie Edwards Paper
b. Rick Reiss/Michael Goodman Paper
c. Gradient Paper (2015)

2. We pointed out that EPA had completed risk assessments for some organophosphates after the epidemiological data were available to them where no FQPA 10x was applied. Here are the specific examples. Here are the specific examples of organophosphates which EPA removed the FQPA 10x and did not reapply until the HED memo issued in 2015:
   a. Bensulide – EPA scoping document in 2008 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data
   b. Phosmet – EPA scoping document in 2009 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data

   Additionally, I also include a letter written in 2013 by Steve Bradbury (then Director of the Office of Pesticide Programs) regarding the use of these same epidemiology data in risk assessments. It is our contention that for EPA to reapply the FQPA 10x to a compound from which the Agency had removed it, they must have reliable and available data. The researchers have not provided the data to the Agency for the epidemiological studies that are the basis for EPA reapplying the FQPA 10x.

3. Ongoing mechanistic data. You mentioned ongoing research– possibly at ORD –to determine if there is some other mode of action occurring at doses lower than those that inhibit cholinesterase that may cause neurodevelopmental effects. Can you please provide more information about what is being done? We would just like to point that previous discussions on this topic often focused on brain rather than RBC and the focus really needs to be on RBC.

Thanks again for agreeing to meet with us on Wednesday- we appreciate it.

Janet

Ex. 6

<Appendix 6_2014 CPFOS HHRA.pdf>

Message

From: Keller, Kaitlin [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=D7A6B15ADFD745C6ADA1C121DEC27AC4-KELLER, KAI]
Sent: 4/11/2018 8:54:18 PM
To: janet collins [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usera98e8fe5]
CC: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]; csmith@gowanco.com [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=85a93dee627e495997f325593ed303eb-csmith@gowa]
Subject: RE: Update Meeting Materials
Attachments: Appendix 6_2014 CPFOS HHRA.pdf

Hi Janet,

As follow-up to the meeting today, I have attached Appendix 6 pulled from the 2014 Chlorpyrifos Human Health Risk Assessment. Section III.B. discusses co-exposure to other environmental contaminants. Also, Russell Carr (out of Mississippi, not North Carolina, but still in the south) is the researcher studying mechanistic effects, specifically related to endocannabinoids.

Thanks,
Kaitlin

Kaitlin Keller, Special Assistant
Office of Chemical Safety and Pollution Prevention
U.S. Environmental Protection Agency
(202) 564-7098

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Monday, April 09, 2018 11:39 AM
To: Keller, Kaitlin <keller.kaitlin@epa.gov>
Cc: Beck, Nancy <Beck.Nancy@epa.gov>; csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Bolen, Derrick <bolen.derrick@epa.gov>
Subject: Re: Update Meeting Materials

Thank you Kaitlin. We can make that work.

On Apr 9, 2018, at 11:15 AM, Keller, Kaitlin <keller.kaitlin@epa.gov> wrote:

> Hi Janet,
>
> Apologies for any confusion. We can do an hour here Wednesday, can you come in 9-10am?
>
> Thank you,
> Kaitlin
>
> Kaitlin Keller, Special Assistant
> Office of Chemical Safety and Pollution Prevention
> U.S. Environmental Protection Agency
> (202) 564-7098

From: Janet Collins [mailto:jcollins@croplifeamerica.org]
Sent: Monday, April 09, 2018 8:13 AM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Cc: csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>; Keller, Kaitlin <keller.kaitlin@epa.gov>; Bolen, Derrick <bolen.derrick@epa.gov>
Subject: Re: Update Meeting Materials

Thanks Nancy- my understanding is that we have one hour.

Please confirm specifically

On Apr 9, 2018, at 7:34 AM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:

> Thanks Janet.
> Should I consider the below information the agenda for our meeting. My understanding is that your group is coming in for 30 minutes on Wednesday.
>
> Thanks,
> Nancy
>
> ---
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> **Ex. 6**
> beck.nancy@epa.gov
>
> **From:** Janet Collins [mailto:jcollins@croplifeamerica.org]
> **Sent:** Monday, April 9, 2018 7:08 AM
> **To:** Beck, Nancy <Beck.Nancy@epa.gov>
> **Cc:** csmith@gowanco.com; Courtney DeMarco <cdemarco@croplifeamerica.org>
> **Subject:** Update Meeting Materials
> **Importance:** High
>
> Nancy- again we really appreciate you joining our Strategic Oversight Council discussion on January 25th. You may recall that we had several items we committed to follow up on for you.
>
> 1. You raised the concept of a 3rd party review of the epidemiological data that is the basis for EPA's HED Memorandum that reapplies an FQPA 10x to all organophosphate risk assessments. We wanted to highlight for you that some 3rd party reviews of those data have been conducted. I have highlighted the summaries of the following papers and provided them in their entirety if you want to review them:
>   a. Debbie Edwards Paper
>   b. Rick Reiss/Michael Goodman Paper
>   c. Gradient Paper (2015)
> 2. We pointed out that EPA had completed risk assessments for some organophosphates after the epidemiological data were available to them where no FQPA 10x was applied. Here are the specific examples. Here are the specific examples of organophosphates which EPA removed the FQPA 10x and did not reapply until the HED memo issued in 2015:

a. Bensulide – EPA scoping document in 2008 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data
b. Phosmet – EPA scoping document in 2009 did not reapply the FQPA 10x based on the epi data despite the Agency being aware of those data

Additionally, I also include a letter written in 2013 by Steve Bradbury (then Director of the Office of Pesticide Programs) regarding the use of these same epidemiology data in risk assessments. It is our contention that for EPA to reapply the FQPA 10x to a compound from which the Agency had removed it, they must have reliable and available data. The researchers have not provided the data to the Agency for the epidemiological studies that are the basis for EPA reapplying the FQPA 10x.

3. Ongoing mechanistic data. You mentioned ongoing research–possibly at ORD – to determine if there is some other mode of action occurring at doses lower than those that inhibit cholinesterase that may cause neurodevelopmental effects. Can you please provide more information about what is being done? We would just like to point that previous discussions on this topic often focused on brain rather than RBC and the focus really needs to be on RBC.

Thanks again for agreeing to meet with us on Wednesday- we appreciate it.

Janet

Ex. 6

Appendix 6. Columbia Center for Children's Environmental Health (CCCEH) Epidemiology Data Acquisition "Raw Data" Request

I. ACTION REQUESTED

To fulfill identified information needs for the purposes of incorporating the Columbia Center for Children's Environmental Health (CCCEH) epidemiology data into the Human Health Risk Assessment (HHRA) for chlorpyrifos, the agency sought to obtain certain "raw data" from CCCEH researchers. Specifically, EPA requested the original analytic data file used to support analyses presented in the peer-reviewed, published epidemiology studies concerning *in utero* chlorpyrifos exposure (V. Rauh et al., 2011; V. A. Rauh et al., 2006; Whyatt et al., 2004). CCCEH researchers did not agree to provide these data, however, the researchers met with EPA and discussed the agency's questions about the data to help determine whether further review of the raw data might assist EPA in resolving uncertainties. As a result of new information gathered through an on-site meeting and other sources, EPA is no longer pursuing the request for the original analytic data file from CCCEH researchers. This memorandum details the new information gained that resolves or renders unobtainable the previously identified information needs.

II. BACKGROUND

EPA considers many different types of scientific information when performing a human health risk assessment (HHRA) of pesticide exposure in the human population. Traditionally, EPA uses toxicology, product and residue chemistry, and industrial hygiene studies as well as measured and modeled human and environmental exposure information to support assessment of environmental risks. In its preparation of the HHRA for chlorpyrifos, the agency has evaluated environmental epidemiology studies of the potential risk of long-term neurodevelopmental effects such as delayed motor skill acquisition or reduced intelligence quotient (IQ) measures among children who experienced pesticide exposure during gestational development. There are three prospective birth cohort studies in the U.S. that examine pesticide exposure (as well as other environmental toxicants) to the pregnant mother and fetus, and then measure neurological and neurodevelopmental performance in children as they grow older. EPA has provided some of

the funding support for each of these studies. Authors hypothesize that *in utero* and early life exposure may influence brain development and effect neurological functioning in children. These studies include the CHAMACOS study in the Salinas Valley, CA, the Mt. Sinai children's environmental health study (Mt. Sinai study), and the Columbia Center for Children's Environmental Health (CCCEH).

The CCCEH study is the only one of the three studies that measures maternal and fetal exposure to chlorpyrifos specifically; the other two cohorts measure exposure to organophosphate pesticides generally. Authors with the CCCEH study reported reduced birth weight and birth length among neonates more highly exposed to chlorpyrifos during gestation (as measured by cord blood concentration of chlorpyrifos) (Whyatt et al., 2004). Similarly, authors observed slower motor skill acquisition and reduced mental capacity among infants who were more highly exposed to the chemical *in utero* (V. A. Rauh et al., 2006). In 2011, authors from all three birth cohort studies concurrently reported evidence of reduced measures of intelligence (Wechslar intelligence scale scores) by increasing *in utero* chlorpyrifos and/or organophosphate exposure (M. F. Bouchard et al., 2011; Engel et al., 2011; V. Rauh et al., 2011).

Given the value of this information to the agency's HHRA for chlorpyrifos, EPA requested the FIFRA SAP to provide external peer review of the strengths and limitations of the epidemiology data for use in the chlorpyrifos HHRA (FIFRA SAP September 2008 and April 2012). The agency identified two major areas in which additional information was needed to fully incorporate these data into the HHRA: additional measures of environmental exposure to chlorpyrifos in the CCCEH cohort to discern whether acetyl cholinesterase inhibition was likely to have occurred in connection with reported adverse outcomes, and also the role of other environmental chemicals (lead, polycyclic aromatic hydrocarbon (PAH), other organophosphate pesticides) in the observed adverse neurological effects reported in relation to *in utero* chlorpyrifos exposure.

To fulfill these information needs for the purposes of incorporating the epidemiology data into the chlorpyrifos HHRA, the agency sought to obtain certain "raw data" from the Columbia Center for Children's Environmental Health (CCCEH) study. Specifically, EPA requested the

original analytic data file used to support analyses presented in the peer-reviewed, published epidemiology studies concerning *in utero* chlorpyrifos exposure (V. Rauh et al., 2011; V. A. Rauh et al., 2006; Whyatt et al., 2004). CCCEH did not agree to provide the data based upon these initial inquiries and they asserted that because EPA did not fund the pesticide exposure component of their cohort study EPA was not legally entitled to review their underlying data. CCCEH did agree, however, to meet and discuss EPA's questions about the data to help determine whether further review of the raw data might assist EPA in resolving uncertainties. As a result on April 15th, 2013, EPA scientists and CCCEH researchers held an all-day meeting at the CCCEH data center (Mailman School of Public Health, New York City, NY) to discuss EPA's information needs and whether acquisition of the full analytic data would be necessary or valuable to EPA's assessment. Addendum 1 delineates the questions EPA posed to CCCEH study staff at this all-day meeting.

III. RESOLUTION OF INFORMATION NEEDS

A. EPIDEMIOLOGY STUDY EXPOSURE CHARACTERIZATION

The primary rationale supporting EPA's request for "raw data" from the CCCEH researchers relates to the agency's need to determine whether the levels of chlorpyrifos exposure in the environment (apartments, apartment building or other outdoor environment, or dietary exposure) of CCCEH study participants were above or below levels that may elicit a greater than 10% inhibition of acetylcholinesterase enzyme levels, the current regulatory endpoint. During the April 2013 meeting, EPA learned that this type of information is neither available nor obtainable. CCCEH researchers estimated relative pesticide exposure using several different exposure methods including 48-hour air sampling with personal monitor, 2-week integrated stationary air monitoring, maternal urinary concentration of TCPy (urinary metabolite of chlopryrifos) during the last trimester of pregnancy, maternal urinary concentration of TCPy at delivery, and umbilical cord blood and meconium at delivery. To determine whether a significant change in acetyl cholinesterase levels may have occurred as a result of actual environmental exposure, temporal concordance between pesticide use and the chlorpyrifos measurement is needed, *i.e.*, exposure estimation at the time of pesticide application is optimal. The CCCEH study design did not incorporate pre- and post-pesticide use/exposure measurement in the study protocol.

Therefore, this information was not collected and is not retrospectively obtainable.

In addition, EPA requested any additional information obtained by researchers as to specific pesticide products used to better understand the pattern and frequency of organophosphate pesticide use among cohort participants. This information was solicited from participants in a written questionnaire administered during a follow-up period (unpublished copy of questionnaire obtained by EPA Oct. 2012). In response to the EPA inquiry, researchers recalled that the Whyatt (2002) publication described the challenges of collecting pesticide product information in etiologic epidemiology studies, and in the on-site meeting in April 2013 confirmed that the information quality in the CCCEH written questionnaire responses is very low. This information was deemed of such poor quality by CCCEH data analysts that the data were not coded or entered into the analytic data file. Therefore, EPA learned that this specific request for "raw data" concerning pesticide product use is not available.

As a surrogate for this information, CCCEH researchers suggested EPA contact the New York City Department of Health to obtain a linked dataset of CCCEH study participant residential address and public housing pesticide usage. The linked dataset provides aggregated pesticide usage data at the cohort participant building-level only. EPA has obtained and reviewed these data (June 2013) and determined that pursuing a data reconstruction exercise is the most appropriate way to estimate environmental pesticide exposure that would have to occur among CCCEH study participants. EPA has conducted such analysis and included it in the revised human health risk assessment.

B. CO-EXPOSURE TO OTHER ENVIRONMENTAL CONTAMINENTS

A second major concern raised by EPA, FIFRA SAP peer reviewers, and public commenters is the ability of the CCCEH study authors to accurately measure and statistically model the relationship between other environmental chemicals (lead and PAH, specifically) or other pesticides (diazinon, propoxur) that may influence fetal brain development and childhood neurodevelopmental performance, and also be related to chlorpyrifos exposure (these are "potentially confounding" exposures). EPA's concern stems from the understanding that if these

other exposures are not sufficiently considered in the epidemiological analysis, then an incorrect inference and conclusion may result (*i.e.*, a potential false positive association). For example, prenatal and early life exposure to lead in the environment has been causally linked to adverse neurodevelopmental outcomes similar to those measured in the CCCEH cohort study including intelligence measures. EPA was concerned about the potential error in the CCCEH study if lead levels were not appropriately considered, *i.e.*, the apparent chlorpyrifos effect on neurodevelopment observed in the study may have been due to the lead exposure.

However, EPA has confirmed with study authors that lead levels and chlorpyrifos levels in cord blood are not statistically associated in this population. Plotting blood lead levels against cord blood chlorpyrifos levels illustrates that the two exposures are extremely weakly (linearly) correlated in this cohort ($\rho<1\%$) (V. A. Rauh et al., 2006). Further, EPA learned from unpublished, supplemental analyses performed by CCCEH researchers upon EPA request that postnatal blood lead levels and prenatal chlorpyrifos levels are also not strongly statistically associated (Andrews, January 21, 2013). This is plausible because of intensive lead abatement programs on-going in New York City during the time period of this study. According to the New York City Department of Health, the number of children with elevated blood lead levels declined 92% between 1995 and 2008.[1] Therefore, because the two exposures are not related, it is not likely that pre- or postnatal blood lead exposure could explain the observed association with chlorpyrifos.

Furthermore, during the April 2013 meeting CCCEH researchers pointed out that based upon available information it appears that lead and chlorpyrifos may affect the brain differently. It is well understood that lead affects the neurodevelopmental sub-domain leading to outward motivation and aggression; while research within the CCCEH cohort indicates chlorpyrifos may affect inward motivation, information processing and organization (V. Rauh et al., 2011; V. A. Rauh et al., 2006; Wright et al., 2008). Additionally, MRI imaging studies of lead affected persons and preliminary brain imaging studies of chlorpyrifos affected persons show different MRI patterns, grey matter as opposed to white matter compositional patterns, respectively (Brubaker, Dietrich, Lanphear, & Cecil, 2010; Brubaker et al., 2009; Cecil et al., 2008; Cecil et

[1] http://www.nyc.gov/html/doh/html/data/stats-childlead.shtml

al., 2011; V. A. Rauh et al., 2012). Therefore, given that neither pre- nor postnatal lead levels and chlorpyrifos levels are not statistically associated with one another in the CCCEH study, and the different ways through which lead and chlorpyrifos appear to influence neurodevelopmental domains EPA concludes that lead exposure did not likely confound (bias or render incorrect) the observed association between chlorpyrifos exposure and neurodevelopment in this study population.

Peer review panelists participating on the April 2012 FIFRA SAP panel identified the concern that authors had not fully considered the long-term effects of polycyclic aromatic hydrocarbon (PAH) exposure, a ubiquitous air pollutant in inner-city areas such as NYC, in the observed association between chlorpyrifos and neurodevelopmental outcomes. Specifically, panelist argued that 'a shift in environmental exposures over time' such that postnatal PAH exposure may have combined with the measured *in utero* pesticide exposure to result in the observed ND outcomes. During the April 2013 meeting, authors clarified that the study design did not include a repeat measure of exposures over time, so an analysis of postnatal PAH exposures is not possible. In the published studies, authors were able to control for the effect of prenatal PAH through statistical adjustment. In addition, authors examined the possible modifying role of prenatal PAH in this epidemiological association and did not observe any evidence of a different risk estimate between chlorpyrifos and ND among those more highly exposed to PAH. Concerning the role of postnatal environmental exposures, CCCEH researchers also stated their belief that their overall study results illustrate that it is gestational exposure, and not early life exposure, that influences neurodevelopment in the study population. They state that the longitudinal analyses of infant and child neurodevelopment in relation to *in utero* chlorpyrifos exposure illustrates a persistent effect of the prenatal environment (M. Bouchard et al., 2003; M. F. Bouchard et al., 2011; Engel et al., 2007; Engel et al., 2011; Eskenazi et al., 2004; Eskenazi et al., 2007; V. Rauh et al., 2011; V. A. Rauh et al., 2006; Whyatt et al., 2004). EPA concluded that CCCEH researchers utilized best practices in statistical analysis of epidemiological data concerning the role of prenatal PAH in neurodevelopmental outcomes, and that a study of repeated, postnatal PAH exposure was beyond the scope of the current CCCEH study, and would require a follow-up study not yet undertaken.

EPA was also interested to learn more about the co-exposure to other organophosphate pesticides among CCCEH study participants. Specifically, EPA as well as external peer review panelists noted the uncertainty as to the degree to which exposure to multiple acetyl cholinesterase inhibiting pesticides exposures over time and/or concurrent in time may have influenced study results. CCCEH researchers agreed that a more clear understanding of the role of mixtures – exposure to multiple OP pesticides overall or concurrent in time – on these neurodevelopmental outcomes is desirable; however they also recognized that the current sample size is too small to perform this type of analysis. To better understand the role of exposure to a mixture of OP pesticides a new cohort study with a larger sample size and different design is required. Therefore, EPA concluded that co-exposure to multiple organophosphate mixtures is not currently obtainable.

For risk characterization purposes, EPA was also interested in understanding the relative contributions of various environmental exposures on ND outcomes, (*e.g.*, PAH, environmental tobacco smoke, chlorpyrifos). Researchers noted that a preliminary indication of the relative contribution of various risk factors for intelligence measures in these cohorts can be seen through examination of supplemental tables published by CCCEH researchers, *i.e.*, the beta-coefficients provided in published supplemental tables provide an indication of the relative contribution of each risk factor (V. Rauh et al., 2011). However, CCCEH researchers indicated that to gain a true reflection the causal model in the population a series of studies in other study populations is needed. EPA and CCCEH researchers agreed that these studies will likely accumulate over time, however they are not currently available.

IV. CONCLUSIONS

In the past, EPA sought to obtain the original analytic data file used to support certain epidemiological analysis of *in utero* exposure to chlorpyrifos and subsequent adverse neurodevelopmental health outcomes in children generated by the Columbia Center for Children's Environmental Health (CCCEH) to support the Human Health Risk Assessment (HHRA) of chlorpyrifos. EPA believed these data were important to both clarify the exposure-response relationship observed in the epidemiology study relative to the current regulatory

endpoint (acetylcholinesterase inhibition), and also to resolve uncertainties regarding study participants co-exposure to other environmental contaminants, among other areas of uncertainties. CCCEH researchers did not agree to provide these data, however, the researchers met with EPA and discussed the agency's questions about the data to help determine whether further review of the raw data might assist EPA in resolving uncertainties. As a result of this meeting and additional discussions with CCCEH staff, EPA concluded that access to the raw data would either not provide answers to EPA's questions or that the information EPA sought could be obtained without analyzing the raw data. Indeed, based on discussions in that meeting as well as further work conducted by agency staff, EPA has gained additional information to better clarify and characterize the major issue areas identified as uncertainties. For these reasons, EPA decided that it would not further pursue its request for the analytic data file from the CCCEH researchers.

Works Cited

Andrews, H. F. (January 21, 2013). [Clarification of Relation between Blood Lead and Cord Blood Levels of Chlorpyrifos in the Columbia Center for Children's Environmental Health (CCCEH) Studies (Electronic mail communication)].

Bouchard, M., Gosselin, N. H., Brunet, R. C., Samuel, O., Dumoulin, M. J., & Carrier, G. (2003). A toxicokinetic model of malathion and its metabolites as a tool to assess human exposure and risk through measurements of urinary biomarkers. *Toxicol Sci, 73*(1), 182-194. doi: 10.1093/toxsci/kfg061

Bouchard, M. F., Chevrier, J., Harley, K. G., Kogut, K., Vedar, M., Calderon, N., . . . Eskenazi, B. (2011). Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. *Environ Health Perspect, 119*(8), 1189-1195. doi: 10.1289/ehp.1003185

Brubaker, C. J., Dietrich, K. N., Lanphear, B. P., & Cecil, K. M. (2010). The influence of age of lead exposure on adult gray matter volume. *Neurotoxicology, 31*(3), 259-266. doi: 10.1016/j.neuro.2010.03.004

Brubaker, C. J., Schmithorst, V. J., Haynes, E. N., Dietrich, K. N., Egelhoff, J. C., Lindquist, D. M., . . . Cecil, K. M. (2009). Altered myelination and axonal integrity in adults with childhood lead exposure: a diffusion tensor imaging study. *Neurotoxicology, 30*(6), 867-875. doi: 10.1016/j.neuro.2009.07.007

Cecil, K. M., Brubaker, C. J., Adler, C. M., Dietrich, K. N., Altaye, M., Egelhoff, J. C., . . . Lanphear, B. P. (2008). Decreased brain volume in adults with childhood lead exposure. *PLoS Med, 5*(5), e112. doi: 10.1371/journal.pmed.0050112

Cecil, K. M., Dietrich, K. N., Altaye, M., Egelhoff, J. C., Lindquist, D. M., Brubaker, C. J., & Lanphear, B. P. (2011). Proton magnetic resonance spectroscopy in adults with childhood lead exposure. *Environ Health Perspect, 119*(3), 403-408. doi: 10.1289/ehp.1002176

Engel, S. M., Berkowitz, G. S., Barr, D. B., Teitelbaum, S. L., Siskind, J., Meisel, S. J., . . . Wolff, M. S. (2007). Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton Neonatal Behavioral Assessment Scale in a multiethnic pregnancy cohort. *Am J Epidemiol, 165*(12), 1397-1404. doi: 10.1093/aje/kwm029

Engel, S. M., Wetmur, J., Chen, J., Zhu, C., Barr, D. B., Canfield, R. L., & Wolff, M. S. (2011). Prenatal exposure to organophosphates, paraoxonase 1, and cognitive development in childhood. *Environ Health Perspect, 119*(8), 1182-1188. doi: 10.1289/ehp.1003183

Eskenazi, B., Harley, K., Bradman, A., Weltzien, E., Jewell, N. A., Barr, D. B., . . . Holland, N. T. (2004). Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. *Environmental Health Perspectives, 112*(10), 1116-1124. doi: 10.1289/ehp.6789

Eskenazi, B., Marks, A. R., Bradman, A., Harley, K., Barr, D. B., Johnson, C., . . . Jewell, N. P. (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. *Environ Health Perspect, 115*(5), 792-798. doi: 10.1289/ehp.9828

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D. B., & Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environ Health Perspect, 119*(8), 1196-1201. doi: 10.1289/ehp.1003160

Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., Barr, D. B., . . . Whyatt, R. W. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. *Pediatrics, 118*(6), e1845-1859. doi: 10.1542/peds.2006-0338

Rauh, V. A., Perera, F. P., Horton, M. K., Whyatt, R. M., Bansal, R., Hao, X. J., . . . Peterson, B. S. (2012). Brain anomalies in children exposed prenatally to a common organophosphate pesticide. *Proceedings of the National Academy of Sciences of the United States of America, 109*(20), 7871-7876. doi: 10.1073/pnas.1203396109

Whyatt, R. M., Camann, D. E., Kinney, P. L., Reyes, A., Ramirez, J., Dietrich, J., . . . Perera, F. P. (2002). Residential pesticide use during pregnancy among a cohort of urban minority women. *Environ Health Perspect, 110*(5), 507-514.

Whyatt, R. M., Rauh, V., Barr, D. B., Camann, D. E., Andrews, H. F., Garfinkel, R., . . . Perera, F. P. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environ Health Perspect, 112*(10), 1125-1132.

Wright, J. P., Dietrich, K. N., Ris, M. D., Hornung, R. W., Wessel, S. D., Lanphear, B. P., . . . Rae, M. N. (2008). Association of prenatal and childhood blood lead concentrations with criminal arrests in early adulthood. *PLoS Med, 5*(5), e101. doi: 10.1371/journal.pmed.0050101

Appendix 6. Addendum 1:

Columbia University Epidemiology Studies

The agency is obligated to review and address peer review comments in support of regulatory decisions. The following is a list of key issues about the epidemiological studies carried out by researchers at Columbia University that were raised in peer review comments. These issues require EPA to have access to the raw data for additional analyses by the agency.

1) **Further analysis of other chemical exposures (*e.g.*, lead, PAHs, other pesticides) to address, if possible, their impact or contribution as modulating factors on the measured outcomes**

- **2012 SAP --** "it should be noted that it cannot be stated that chlorpyrifos is the sole contributor to the observed outcomes."
- **2012 SAP --** "In an earlier examination of the same cohort, Perera *et al.*(2009) reported an association between a decrease in full-scale IQ and verbal IQ in 5year-olds with prenatal polycyclic aromatic hydrocarbons (PAH) exposure rather than chlorpyrifos, thus, raising an issue of the shift in chemical exposure association with increase in age. In each of these analyses, statistical modeling showed that the exposures were independently associated with IQ, and no significant interaction was observed with the other chemical. While this is a statistically sound approach to determine independent responses, panel members noted that it is very difficult to identify the independent physiological effects of a single chemical in this type of multi-chemical exposure scenario."
- **2012 Federal Peer Review --** "even low levels of lead can impact neurodevelopment, and even that the observed neurobehavioral deficits are more pronounced at lower blood lead levels when compared with higher blood lead levels".
- **2008 SAP --** "In order to eliminate the possible causes of neurodevelopmental effects by other pesticides in the Columbia study, it is suggested that EPA should repeat the pre-post residential cancellation analysis done for chlorpyrifos using other pesticide measurements, such as malathion diacid (MDA), a specific metabolite of malathion. The outcomes from those additional analyses will either confirm or reject EPA's preliminary conclusion that chlorpyrifos is likely to play a role in the neurodevelopmental outcomes."
- **2008 SAP --** ""It would be useful to examine the results of a statistical analysis that includes all three AChE-inhibiting insecticides in the analysis model as dichotomous variables (above or below LOD) in combination with continuous measurements for these variables. This type of analysis would likely not change

the results, but it could be helpful in illustrating threshold or dose response effects."

2) **Further analysis and information to address and, if possible, better characterize uncertainty around outcome measures on learning/memory/IQ**

- **2012 SAP**-- Alternative considerations for non-quantified samples: "little use was made of techniques to integrate non-quantified samples into the statistical test.... Various methods were reviewed by the July 2010 SAP that can be applied to either normally or lognormally distributed data that include a significant (even a majority) of non-detectable sample . . . . Specifically, the use of 'probability plots' was described that can yield an estimate of the geometric mean of the distribution [GM], the geometric standard deviation [GSD], and corresponding percentiles."
- **Federal Peer Review --** "There is a scatterplot showing the raw scores for overall IQ and for each of the subtests, but it is not possible to obtain the necessary information to compare the distributions of these scores with the norms for the test or with any other study sample. Ideally, the means and standard deviations for these scores should be presented for either a non-exposed or a non-exposed combined with low exposed group and these should be compared to a moderate or high-exposed group as was done for the BSID-II in the Rauh et al., 2006 paper. Here the uncertainties stem from the assumptions that are made when regression analyses are performed. The main issue here is that outliers can greatly influence the slope of the function."
- **Federal Peer Review--**A between group analysis using inferential statistics, as was done for the Bayley Scales of Infant Development II in the Rauh et al., 2006 paper, should be performed on each variable in both studies (i.e., the Child Behavior Checklist in Rauh et al., 2006, and the full scale IQ and subscales for the WISC-IV in the Rauh et al., 2011 study). This would be the most direct and least problematic method for determining whether exposure to chlorpyrifos resulted in significant decreases in IQ or significant increases in behavioral problems "..... no information was provided regarding the qualifications of the individuals who administered and scored the tests. "

3) **Further analysis to assess, if possible, whether individual cohort members had the potential for exposure to chlorpyrifos and/or other acetylcholinesterase (AChE) inhibiting pesticides (e.g., diazinon, propoxur), prenatally and /or postnatally, at levels leading to greater than 10% AChE inhibition (the level used to derive the regulatory point of departure).**

- **2012 SAP--** recommended conducting a dose reconstruction analysis—"data on the concentration of chlorpyrifos in various media (*i.e.* house dust, air and water) while market basket data exists on the concentration of chlorpyrifos on food. These data provide the main tools for developing an effective exposure assessment and a subsequent reconstruction of potential dose." The agency has begun such analysis but the current draft analysis is limited without data on the exposure information relevant to individual women such that environmental chlorpyrifos exposure can then be linked to measures of blood chlorpyrifos.
- **2012 SAP--** recommended the agency consider issues related to multiple chemical exposure (i.e., mixtures) to chlorpyrifos and other key AChE inhibiting pesticides identified by the Columbia University studies (diazinon, propoxur). Assumptions of co-exposure will likely be grossly overestimated without access to the raw data; such raw data may enable the agency to evaluate actual co-exposure information for individuals from air monitoring samples and blood samples.

Message

From: Jay Vroom [JVroom@croplifeamerica.org]
Sent: 9/7/2017 7:34:39 PM
To: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
Subject: In advance of CLA Meeting . . .

Good afternoon, Nancy,

Just a quick note in advance of our 10:15a meeting tomorrow, to let you know that I'm bringing two colleagues with me, Janet Collins and Rachel Lattimore, as well as representatives from 3 of our member companies:

- Cindy Smith, Gowan
- Michael Parrish, Monsanto
- Eric Tamichi, Valent

I understand you are meeting at 11:00a tomorrow with RISE, which is closely affiliated with CLA, and I plan to stay on for that meeting, too!

See you in the morning.
Jay

Jay Vroom
President & CEO
CropLife America
1156 15th Street, NW
Suite 400
Washington, DC 20005

Ex. 6

Fax (202) 466-5832
Email vroom@croplifeamerica.org
Executive Assistant Mary Jo Tomalewski (mjtomalewski@croplifeamerica.org, Ex. 6
Web www.croplifeamerica.org

Message

From: Jay Vroom [JVroom@croplifeamerica.org]
Sent: 8/26/2017 4:27:24 AM
To: live.com#michael.parrish@monsanto.com [michael.parrish@monsanto.com]; Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
Subject: Conversation support on dicamba issues

Hello Mike,

I had a phone call with Nancy yesterday from a fairly noisy spot in an airport so our conn cation was not the best-- and on the topic of the evolving scenario of dicamba issues this year I think I understood her to say she was waiting on an attempted connection with one of your St Louis colleagues. I told Nancy that you could
Likely help facilitate any such conversations and that I'd send this email to extend that idea to you both!

Jay

Sent from my iPhone

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/28/2017 10:15:25 PM
To: PARRISH, MICHAEL [AG/1920] [michael.parrish@monsanto.com]; Jay Vroom [JVroom@croplifeamerica.org]
Subject: RE: Conversation support on dicamba issues

Thanks Mike.
I actually had a meeting with some Monsanto folks this morning so I think we are in good shape.

Regards,
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
P: 202-564-1273
M: Ex. 6
beck.nancy@epa.gov

-----Original Message-----
From: PARRISH, MICHAEL [AG/1920] [mailto:michael.parrish@monsanto.com]
Sent: Monday, August 28, 2017 10:14 AM
To: Jay Vroom <JVroom@croplifeamerica.org>; Beck, Nancy <Beck.Nancy@epa.gov>
Subject: RE: Conversation support on dicamba issues

Thanks Jay.

Nancy - My contact info is below as well. Let me know when you would like to connect.

Mike

Mike Parrish
Monsanto | N.A. Corporate Engagement Lead Mobile Ex. 6 Discover Monsanto - www.discover.monsanto.com America's Farmers Grow America - Learn more at www.americasfarmers.com

-----Original Message-----
From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Saturday, August 26, 2017 12:27 AM
To: PARRISH, MICHAEL [AG/1920] <michael.parrish@monsanto.com>; Nancy Beck (PhD, DABT) <beck.nancy@epa.gov>
Subject: Conversation support on dicamba issues

Hello Mike,

I had a phone call with Nancy yesterday from a fairly noisy spot in an airport so our conn cation was not the best-- and on the topic of the evolving scenario of dicamba issues this year I think I understood her to say she was waiting on an attempted connection with one of your St Louis colleagues. I told Nancy that you could Likely help facilitate any such conversations and that I'd send this email to extend that idea to you both!

Jay

Sent from my iPhone

Message

From: PARRISH, MICHAEL [AG/1920] [michael.parrish@monsanto.com]
Sent: 8/29/2017 6:34:42 PM
To: Beck, Nancy [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=168ecb5184ac44de95a913297f353745-Beck, Nancy]
CC: Jay Vroom [JVroom@croplifeamerica.org]
Subject: Re: Conversation support on dicamba issues

Thanks Nancy. I connected with Ty and Tom and realized you were meeting yesterday.

Thanks,
Mike

> On Aug 28, 2017, at 5:15 PM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:
>
> Thanks Mike.
> I actually had a meeting with some Monsanto folks this morning so I think we are in good shape.
>
> Regards,
> Nancy
>
> _____
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
> P: 202-564-1273
> M: Ex. 6
> beck.nancy@epa.gov
>
> -----Original Message-----
> From: PARRISH, MICHAEL [AG/1920] [mailto:michael.parrish@monsanto.com]
> Sent: Monday, August 28, 2017 10:14 AM
> To: Jay Vroom <JVroom@croplifeamerica.org>; Beck, Nancy <Beck.Nancy@epa.gov>
> Subject: RE: Conversation support on dicamba issues
>
> Thanks Jay.
>
> Nancy - My contact info is below as well. Let me know when you would like to connect.
>
> Mike
>
>
> Mike Parrish
> Monsanto | N.A. Corporate Engagement Lead Mobile Ex. 6 Discover Monsanto - www.discover.monsanto.com America's Farmers Grow America - Learn more at www.americasfarmers.com
>
>
>
> -----Original Message-----
> From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
> Sent: Saturday, August 26, 2017 12:27 AM
> To: PARRISH, MICHAEL [AG/1920] <michael.parrish@monsanto.com>; Nancy Beck (PhD, DABT) <beck.nancy@epa.gov>
> Subject: Conversation support on dicamba issues
>
> Hello Mike,
>
> I had a phone call with Nancy yesterday from a fairly noisy spot in an airport so our conn cation was not the best-- and on the topic of the evolving scenario of dicamba issues this year I think I understood her to say she was waiting on an attempted connection with one of your St Louis colleagues. I told Nancy that you could Likely help facilitate any such conversations and that I'd send this email to extend that idea to you both!
>
> Jay
>
> Sent from my iPhone
> This email and any attachments were sent from a Monsanto email account and may contain confidential and/or privileged information. If you are not the intended recipient, please contact the sender and delete this email and any attachments immediately. Any unauthorized use, including disclosing, printing, storing, copying or distributing this email, is prohibited. All emails and attachments sent to or from Monsanto email accounts may be subject to monitoring, reading, and archiving by Monsanto, including its affiliates and subsidiaries, as permitted by applicable law. Thank you.

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/15/2018 9:28:20 PM
To: janet collins [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=usera98e8fe5]
Subject: Re: If tonight doesn't work

8am or 10:30 should work for me. With 8am being more reliable. Thanks

****

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator
Office of Chemical Safety and Pollution Prevention
Ex. 6
beck.nancy@epa.gov

On Aug 15, 2018, at 5:27 PM, Janet Collins <jcollins@croplifeamerica.org> wrote:

> Let's talk tomorrow!
>
> Sent from my iPhone

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/4/2017 10:12:03 PM
To: Jay Vroom [JVroom@croplifeamerica.org]
Subject: RE: I have to board my flight now

You too.

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
beck.nancy@epa.gov

-----Original Message-----
From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Friday, August 4, 2017 6:06 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Subject: Re: I have to board my flight now

Ok thanks-- I'll try in that transit window for you-- it's probably overlapping when I'll be at Dulles in advance of my flight departure. Have a great weekend..

Jay

Sent from my iPhone

> On Aug 4, 2017, at 6:01 PM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:
>
> Wow-- that should be a great trip.
> I land somewhere around 12pm PST and have events 4:30-9pm PST Monday night. So perhaps there is a small window while I think I will be in transit by car. You can try my cellphone.
>
> Regards,
> Nancy
>
>
> Nancy B. Beck, Ph.D., DABT
> Deputy Assistant Administrator, OCSPP
>
> Ex. 6
>
> beck.nancy@epa.gov
>
>
> -----Original Message-----
> From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
> Sent: Friday, August 4, 2017 5:49 PM
> To: Beck, Nancy <Beck.Nancy@epa.gov>
> Subject: Re: I have to board my flight now
>
> Nancy,
> I am flying to Africa Monday night and won't return to the USA until August 26. Any chance you'd have a window Monday we might connect?
> Thanks
> Jay
> Sent from my iPhone
>
>> On Aug 4, 2017, at 5:34 PM, Beck, Nancy <Beck.Nancy@epa.gov> wrote:
>>
>> Jay,
>> I'm stuck in another meeting. Sorry we didn't connect. I will be back in the office on the 14th.
>>
>> Regards,
>> Nancy
>>
>>
>> Nancy B. Beck, Ph.D., DABT
>> Deputy Assistant Administrator, OCSPP
>>
>> Ex. 6
>>
>> beck.nancy@epa.gov

>>
>> -----Original Message-----
>> From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
>> Sent: Friday, August 4, 2017 5:30 PM
>> To: Beck, Nancy <Beck.Nancy@epa.gov>
>> Subject: I have to board my flight now
>>
>> Let me know when we might try to connect again!
>>
>> Thanks!
>>
>> Sent from my iPhone

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/4/2017 9:34:39 PM
To: Jay Vroom [JVroom@croplifeamerica.org]
Subject: RE: I have to board my flight now

Jay,
I'm stuck in another meeting. Sorry we didn't connect. I will be back in the office on the 14th.

Regards,
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
beck.nancy@epa.gov

-----Original Message-----
From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Friday, August 4, 2017 5:30 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Subject: I have to board my flight now

Let me know when we might try to connect again!

Thanks!

Sent from my iPhone

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 9/10/2017 11:53:15 PM
To: Jay Vroom [JVroom@croplifeamerica.org]; Jakob, Avivah [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=ca1aecd941984ff2939fe77425b0e2f3-Jakob, Avivah]; Keigwin, Richard [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=151baabb6a2246a3a312f12a706c0a05-Richard P Keigwin Jr]
Subject: RE: Thank you for meeting CropLife and RISE today

Jay,
Thanks for coming in. It was particularly helpful to learn about the mosquito spraying delays—thank you for sharing that information with us. We made sure that all the right parties, including the emergency operations center, were aware of what happened, so hopefully we wont see a repeat.

As always, the weekend is too short. I hope you had a good one.
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
Beck.Nancy@epa.gov

From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Friday, September 8, 2017 6:59 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>; Jakob, Avivah <Jakob.Avivah@epa.gov>; Keigwin, Richard <Keigwin.Richard@epa.gov>
Subject: Thank you for meeting CropLife and RISE today

Dear Nancy, Rick and Avivah,

I am sure it's been a long day over there in your shop—as it has been here. Just a quick note to say thanks for making time for the back to back meetings today with CLA and RISE. We deeply value your time and attention to important dialogue around some very critical issues. We owe you essential follow up on several points and will be getting back to you soon.

Meantime, I hope you all have started very peaceful weekends by now!

Jay

Jay Vroom
President & CEO
CropLife America
1156 15th Street, NW
Suite 400
Washington, DC 20005
Ex. 6
Vroom@croplifeamerica.org
www.croplifeamerica.org

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/3/2017 10:12:31 PM
To: Jay Vroom [JVroom@croplifeamerica.org]
Subject: RE: Just left you a voicemail in your office phone

Hi Jay,
Happy to chat. I've got meetings til about 3pm tomorrow but should be available afterwards.
Next week I will be doing the Specialty Crop Tours in California all week. I'm told the schedule is pretty packed so I'm not sure what amount of free time I will have. However, you can always try my cell (number below)
If tomorrow doesn't work, I'll be back in the office on the 14th.

Regards,
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
beck.nancy@epa.gov

-----Original Message-----
From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Thursday, August 3, 2017 4:18 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Subject: Just left you a voicemail in your office phone

Hi Nancy,

Hoping we might have a chat by phone through tomorrow or maybe even in person early next Monday. Let me know if you might have time for a call tomorrow or Monday?

Jay

Jay Vroom
CropLife America
Ex. 6

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 7/19/2017 6:00:15 PM
To: Jay Vroom [JVroom@croplifeamerica.org]
CC: Avivah Jakob (Jakob.Avivah@epa.gov) [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=ca1aecd941984ff2939fe77425b0e2f3-Jakob, Avivah]
Subject: RE: Good meeting today

Jay,
Yes, the meeting was indeed helpful for me. In an ideal world, next time the discussions will happen well in advance of our having to make a decision-- although I'm beginning to recognize that 24 hours, in some cases, is not too bad.

Unfortunately, I don't think I will be able to make it next week. I'm not sure what happened to the invite, but my calendar is already packed both days and it would be extremely hard for me to step away. I would welcome an invitation when your members are back in town again in the future and perhaps by then we can talk about some substantive progress rather than just our plans and intentions. If we give Avivah enough notice, she should be able to work with our schedulers to make it happen.

Regards,
Nancy

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator, OCSPP
Ex. 6
beck.nancy@epa.gov

-----Original Message-----
From: Jay Vroom [mailto:JVroom@croplifeamerica.org]
Sent: Monday, July 17, 2017 8:40 PM
To: Beck, Nancy <Beck.Nancy@epa.gov>
Subject: Good meeting today

Nancy

Thanks for your time this afternoon. Please do not hesitate call if additional questions arise.

I also sincerely hope you can make time to meet with CropLife's Strategic Oversight Council next week (Tues July 25 and Wed July 26)-- per our earlier transmitted invite. Our member leaders will gain much from a good conversation with you.

Jay

Jay Vroom
CropLife America
Ex. 6

Sent from my iPhone

Message

From: Beck, Nancy [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=168ECB5184AC44DE95A913297F353745-BECK, NANCY]
Sent: 8/6/2018 4:24:18 PM
To: Jay Vroom [JVroom@croplifeamerica.org]
CC: Keigwin, Richard [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=151baabb6a2246a3a312f12a706c0a05-Richard P Keigwin Jr]
Subject: Re: Andrew Wheeler--Meeting today with ESA letter signatories?

Jay,
I will be attending.

Regards,
Nancy

****

Nancy B. Beck, Ph.D., DABT
Deputy Assistant Administrator
Office of Chemical Safety and Pollution Prevention
Ex. 6
beck.nancy@epa.gov

On Aug 6, 2018, at 9:55 AM, Jay Vroom <JVroom@croplifeamerica.org> wrote:

> Hi Nancy, Rick--
>
> Are either or both of you attending this meeting with the Acting Administrator today at 4pm<
>
> Jay
>
> *Jay Vroom*
> President & CEO
> CropLife America
> 1156 15th Street, NW
> Suite 400
> Washington, DC 20005
> Ex. 6
> **Email**: vroom@croplifeamerica.org
> **Executive Assistant**: Mary Jo Tomalewski (202.872.3849, mjtomalewski@croplifeamerica.org)

Message

From: Ethan Mathews [emathews@croplifeamerica.org]
Sent: 1/31/2018 10:33:51 PM
To: Bennett, Tate [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=1fa92542f7ca4d01973b18b2f11b9141-Bennett, El]
CC: Sands, Jeffrey [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=b2aa28629ade4afb8d5ec12a8aaaba54-Sands, Jeff]
Subject: Re: Executed MOA

Thank you!!

Ethan Mathews
Director of Government Affairs
CropLife America
emathews@croplifeamerica.org
Ex. 6 (o)
(m)

On Jan 31, 2018, at 5:32 PM, Bennett, Tate <Bennett.Tate@epa.gov> wrote:

From: Bennett, Tate
Sent: Wednesday, January 31, 2018 5:13 PM
To: 'Beau Greenwood' <BGreenwood@croplifeamerica.org>; 'ematthews@croplifeamerica.org' <ematthews@croplifeamerica.org>
Subject: Executed MOA

<Executed ESA-FIFRA MOA 1.31.18.pdf>

Message

From: Greenwalt, Sarah [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=6C13775B8F424E90802669B87B135024-GREENWALT,]
Sent: 3/5/2018 1:13:12 PM
To: Bennett, Tate [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=1fa92542f7ca4d01973b18b2f11b9141-Bennett, El]; Mary Jo Tomalewski [mjtomalewski@croplifeamerica.org]
Subject: RE: Invitation to Meet

My apologies, but I have a meeting with the Administrator at that time.

Sarah A. Greenwalt

U.S. Environmental Protection Agency
Work: 202-564-1722|Cell: Ex. 6
Greenwalt.Sarah@epa.gov

-----Original Message-----
From: Bennett, Tate
Sent: Friday, March 2, 2018 12:35 PM
To: Mary Jo Tomalewski <mjtomalewski@croplifeamerica.org>
Cc: Greenwalt, Sarah <greenwalt.sarah@epa.gov>
Subject: Re: Invitation to Meet

I am available should our standing 8:30 run on time.

> On Mar 2, 2018, at 12:17 PM, Mary Jo Tomalewski <mjtomalewski@croplifeamerica.org> wrote:
>
> Good afternoon,
>
> Jay Vroom from CropLife America asked me to reach out to you to invite you to an hour-long meeting that we are having on Tuesday, March 6 at 9 AM, with Henry Darwin in his offices. A group of our Board of directors and other industry leaders are in town for CLA winter board meeting and they want to meet to discuss a number of issues and EPA processes.
>
> If you are available we would be delighted if you would join us.
>
> MJ
>
> Sent from my iPhone~Please excuse any typos!

Message

From: Beau Greenwood [BGreenwood@croplifeamerica.org]
Sent: 1/31/2018 10:33:10 PM
To: Bennett, Tate [/o=ExchangeLabs/ou=Exchange Administrative Group (FYDIBOHF23SPDLT)/cn=Recipients/cn=1fa92542f7ca4d01973b18b2f11b9141-Bennett, El]
CC: ematthews@croplifeamerica.org
Subject: Re: Executed MOA

Thanks Tate!!

Beau.

On Jan 31, 2018, at 5:13 PM, Bennett, Tate <Bennett.Tate@epa.gov> wrote:

<Executed ESA-FIFRA MOA 1.31.18.pdf>

Delivery Report

From: postmaster@croplifeamerica.org [postmaster@croplifeamerica.org]
Sent: 2/2/2018 9:37:05 PM
To: ematthews@croplifeamerica.org
Subject: Undeliverable: Executed MOA
Attachments: Executed MOA

Your message
To: 'Beau Greenwood'; ematthews@croplifeamerica.org
Subject: Executed MOA
Sent: 1/31/2018 10:13:24 PM

Delivery has failed to these recipients or groups:

ematthews@croplifeamerica.org (ematthews@croplifeamerica.org)
Your message couldn't be delivered. Despite repeated attempts to contact the recipient's email system it didn't respond.

Contact the recipient by some other means (by phone, for example) and ask them to tell their email admin that it appears that their email system isn't accepting connection requests from your email system. Give them the error details shown below. It's likely that the recipient's email admin is the only one who can fix this problem.

For more information and tips to fix this issue see this article: https://go.microsoft.com/fwlink/?LinkId=389361.

Diagnostic information for administrators:

Generating server: SN4PR0501MB3712.namprd05.prod.outlook.com
Receiving server: SN4PR0501MB3712.namprd05.prod.outlook.com
Total retry attempts: 50

ematthews@croplifeamerica.org
2/2/2018 9:37:05 PM - Server at SN4PR0501MB3712.namprd05.prod.outlook.com returned '550 5.4.300 Message expired'

Original message headers:

```
Received: from BN6PR05CA0034.namprd05.prod.outlook.com (2603:10b6:405:39::47)
 by SN4PR0501MB3712.namprd05.prod.outlook.com (2603:10b6:803:48::22) with
 Microsoft SMTP Server (version=TLS1_2,
 cipher=TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA384_P256) id 15.20.464.6; Wed, 31
 Jan 2018 22:13:38 +0000
Received: from DM3NAM05FT053.eop-nam05.prod.protection.outlook.com
```